COA Tracking ID: C2022353
BLA Number: 125771

Niloofar M. Kennedy -S
Digitally signed by Niloofar M. Kennedy -S
Date: 2023.02.21 15:34:00 -05'00'

CLINICAL OUTCOME ASSESSMENT (COA) CONSULT REVIEW

COA Tracking ID:	C2022353
BLA#:	125771
Applicant:	Bioverativ Therapeutics, Inc.
Established Name/Trade Name:	Altuviiio (efanesoctocog alfa, rFVIIIFc-VWF-XTEN)
Indication:	Treatment of adults and children with Hemophilia A (congenital Factor VIII deficiency) for: (1) Routine prophylaxis to reduce the frequency of bleeding episodes; (2) On-demand treatment and control of bleeding episodes; (3) Perioperative management of bleeding. Please check all that apply: ☒Rare Disease/Orphan Designation ☒Pediatrics
PDUFA Goal Date:	PDUFA Goal Date
Review Division:	Office of Tissues and Advanced Therapies (OTAT), Center for Biologics Evaluation and Research (CBER)
Clinical Reviewer/Medical Officer:	Megha Kaushal, MD, MSc
Regulatory Project Manager:	Niloofar Kennedy
COA Reviewer:	Sarah Stothers, RN, MSN, MPH
COA TL:	David Reasner, Ph.D.
Patient Focus Statistical Support (PFSS) Team Leader, OTS/OB/DBIII	Lili Garrard, PhD
COA Division Director:	David Reasner, Ph.D.
Instruments reviewed:	1. Hemophilia Quality of Life Questionnaire for Adults (Haem-A-QoL) ☒ Patient-reported outcome (PRO)
	2. Patient-Reported Outcomes Measurement Information System (PROMIS) Pain Intensity 3a ☒ Patient-reported outcome (PRO)

Table of Contents

1. EXECUTIVE SUMMARY

This Clinical Outcome Assessment (COA) consult review was for BLA 125771 in which the applicant seeks approval of Altuviiio (efanesoctocog alfa, rFVIIIFc-VWF-XTEN) for the treatment of adults and children with Hemophilia A (congenital Factor VIII deficiency) for: (1) routine prophylaxis to reduce the frequency of bleeding episodes; (2) on-demand treatment and control of bleeding episodes; (3) perioperative management of bleeding.

The Office of Tissues and Advanced Therapies (OTAT), Center for Biologics Evaluation and Research (CBER) sought COA input for two patient-reported outcome (PRO) measures (b) (4) .

Table 1. Secondary Efficacy PRO Endpoints (b) (4)

Instrument Name	COA Type	Concept(s)	Endpoint (Study Name)	Copy of Instrument
Hemophilia-specific Quality of Life (Haem-A-QoL; Questionnaire for adults) Physical Health (PH) subscale	PRO	Physical health impairments attributed to hemophilia	Secondary (XTEND-1)	See Appendix A
Patient-Reported Outcomes Measurement Information System (PROMIS) Pain Intensity 3a Item 1 (Worst Pain)	PRO	Pain intensity	Secondary (XTEND-1)	See Appendix B

The PRO-based secondary efficacy endpoint results (Haem-A-QoL PH and PROMIS Pain Intensity 3a Item 1 [Worst Pain]) from Study EFC 16293 (henceforward referred to as the XTEND-1 study throughout this review document**)** were challenging to interpret (b) (4) due to the following:

a. The open label nature of the XTEND-1 study design may have led to biased responses for the PRO measures (i.e., patients' knowledge of treatment assignment is likely to influence how they report information on the PRO) and, subsequently, to biased estimates of treatment effect.
b. Exit interviews used to support clinically meaningful within-patient change for the PRO-based secondary efficacy endpoints were optional (i.e., convenience sample) and the sample selection may have led to biased responses.
c. Quantitative anchor-based analyses and results from the exit interviews suggested that "no change" as measured by the PRO instruments may also be considered important and meaningful to patients; (b) (4)

d. Lack of sensitivity due to a floor effect (i.e., patients do not have sufficient symptom/functional impairment) was observed on the Haem-A-QoL PH subscale and PROMIS Pain Intensity 3a (Worst Pain) item scores for Arm A at baseline, and a small magnitude of change in the PRO secondary endpoint scores at Week 52.
e. The range of clinically meaningful within-patient change thresholds was derived, in part, by distribution-based methods. Distribution-based approaches are inappropriate as a primary method for the evaluation of clinically meaningful within-patient change, as they do not account for the patient voice and/or perspective.

2 REVIEW CONCLUSIONS

Haemophilia Quality of Life Questionnaire for Adults (Haem-A-QoL) – Physical Health Subscale

The Haem-A-QoL was reviewed for content validity, other measurement properties, and meaningful within-patient change in patients with Hemophilia A. (b) (4) :

- **Issue 1:** Potential floor effect due to relatively low item scores at baseline and raw change scores at week 52.
- **Issue 2:** "No change," or "staying the same," may be just as meaningful to patients as an 'improvement" as evidenced by the quantitative anchor-based analyses for clinically meaningful change using raw scores and results from the optional exit interviews.
- **Issue 3:** A clinically meaningful change threshold (e.g., upper bound of -4.8) was calculated using a distribution-based approach. Distribution-based approaches are inappropriate as the primary method for calculating clinically meaningful within patient change, as they do not account for the patient voice and/or perspective.

Patient-Reported Outcomes Measurement Information System Short Form v 1.0 (PROMIS-SF v 1.0) Pain Intensity 3a

The PROMIS Pain Intensity 3a (worst pain) was reviewed for content validity, other measurement properties, and meaningful within-patient change in the Applicant's study population. (b) (4) :

- **Issue 1:** Potential floor effects due to relatively low item scores at baseline and change scores at week 52.
- **Issue 2:** "No change" may be just as meaningful to patients as an 'improvement" as evidenced by the anchor-based analyses for clinically meaningful change using raw scores and results from the optional exit interviews.
- **Issue 3:** Very weak correlation (Pearson's correlation ≤ 0.17) between change at baseline to week 52 on the PROMIS pain item and the primary endpoint, annualized bleeding rate (ABR) were observed, suggesting that change in pain intensity is disassociated from other study measures.
- **Issue 4:** A clinically meaningful change threshold (e.g., upper bound of -0.2) was calculated using a distribution-based method. Distribution-based approaches are inappropriate as the primary methods for calculating clinically meaningful within patient change, as they do not account for the patient voice and/or perspective.

3 RECOMMENDATIONS FOR FUTURE STUDIES

For future clinical trials in this indication, it may be helpful to explore the following methods:

- **the use of a comparator arm.**
- **The use of stratification, random assignment, and other methods in recruiting for exit interviews which may help to reduce biased responses.**
- **evaluating whether "No change," or "No worsening" is meaningful and relevant to patients.**

4 BACKGROUND AND CORRESPONDENCE ON CLINICAL OUTCOME ASSESSMENT(S)

Regulatory Background:

- For Regulatory and Background Materials reviewed, refer to Section 7.3.

Disease Background:

Hemophilia A is a rare (estimated incidence 1 in 500 male births) X-chromosome-linked bleeding disorder that predominantly affects males. It is characterized by the deficiency of functional coagulation factor (FVIII) leading to potentially life-threatening bleeding in response to trauma and recurrent bleed in major joints, soft tissues, and muscle. Currently, there is no cure for Hemophilia A.

- Symptoms include joint pain, bleeding with minor or no trauma, bruising, joint deformation, hemorrhage, and/or thromboembolism.

Investigational Product:

Altuviiio (efanesoctocog alfa, rFVIIIFc-VWF-XTEN) is a recombinant fusion protein consisting of single-chain FVIII, the Fc domain of human immunoglobulin G1 (IgG1), the FVIII-binding D′D3 domain of von Willebrand factor (VWF), and 2 XTEN linkers. It is the first rFVIII protein specifically designed to be independent of endogenous VWF in order to extend the half-life of the FVIII molecule in plasma. A prolonged half-life has been demonstrated to show efficacy and safety in preventing bleeding episodes when used prophylactically, in treating bleeding episodes, and in maintaining homeostasis in perioperative settings.

(NOTE: Altuviiio (efanesoctocog alfa, rFVIIIFc-VWF-XTEN) is referred to as BIVV001 in the Applicant's clinical studies.)

Materials reviewed:

(b) (4) the applicant submitted a(n):

- Clinical Study Report Phase 3.
- Annotated Package insert.
- Patient-Reported Outcome (PRO) Evidence Dossier (b) (4) .
- Response to FDA Information Request [dated 16 November 2022].

5 CLINICAL OUTCOME ASSESSMENT REVIEW

5.1 Clinical Trial Design

The XTEND-1 was a phase 3 multicenter, multinational, non-randomized, open-label study to evaluate the safety, efficacy and pharmacokinetics of BIVV001 in previously treated patients with severe hemophilia A (defined as <1 IU/dL [<1%] endogenous FVIII or a documented genotype known to produce severe hemophilia A).

Approximately 159 participants aged ≥ 12 years were enrolled into two treatment arms (Arm A and B, as listed below). All participants (except a subgroup of Arm A [n=13] who were required to estimate the terminal half-life) underwent a washout period (at least 4 to 5 days, depending on current therapy) prior to the first dose of BIVV001 (Baseline).

- **Arm A:** Included participants who were on a prophylaxis treatment regimen with FVIII prior to the study. One hundred twenty-four (n=124) participants received BIVV001 at a dose of 50 IU/kg IV once weekly (QW) on a prophylaxis treatment regimen for up to 52 weeks.
- **Arm B:** Included participants who were on an on-demand treatment regimen prior to the study. Twenty-six (n=26) participants received BIVV001 (50 IU/kg) on demand for 26 weeks and then switched to prophylactic QW dose (50 IU/kg) of BIVV001 QW for another 26 weeks.

NOTE: Only Arm A was used to evaluate efficacy in the pivotal Phase 3 XTEND-1 study.

A complete list of the inclusion and exclusion criteria is summarized in Section 7.1 of Protocol G1T28-207 version 3.0 (located in Appendix 1 of the Type C Meeting Package Materials [dated 02 November 2022]).

Primary Endpoint

- Mean Annual bleeding rate (ABR) in Arm A (see Figure 1)

Figure 1. Annual Bleeding Rate (ABR) Calculation

$$ABR = \frac{Number\ of\ treated\ bleeding\ episodes\ during\ the\ efficacy\ period}{Total\ number\ of\ days\ during\ the\ efficacy\ period} \times 365.25$$

Secondary COA-based Endpoint(s)

- Change in Haem-A-QoL (≥ 17 years old) total physical health PH) subscale score measures from Baseline to Week 52 in Arm A
- Change in PROMIS Pain Intensity 3a from Baseline to Week 52 in Arm A

After completion of the XTEND-1 study, all participants were invited to continue in an open-label long-term extension study (LTS16294, henceforward referred to as XTEND-ed).

Reviewer's comment(s):

- ***EXTEND-1 was an open-label study. In general, an open-label trial design limits interpretability of PRO data because patients' knowledge of treatment assignment may***

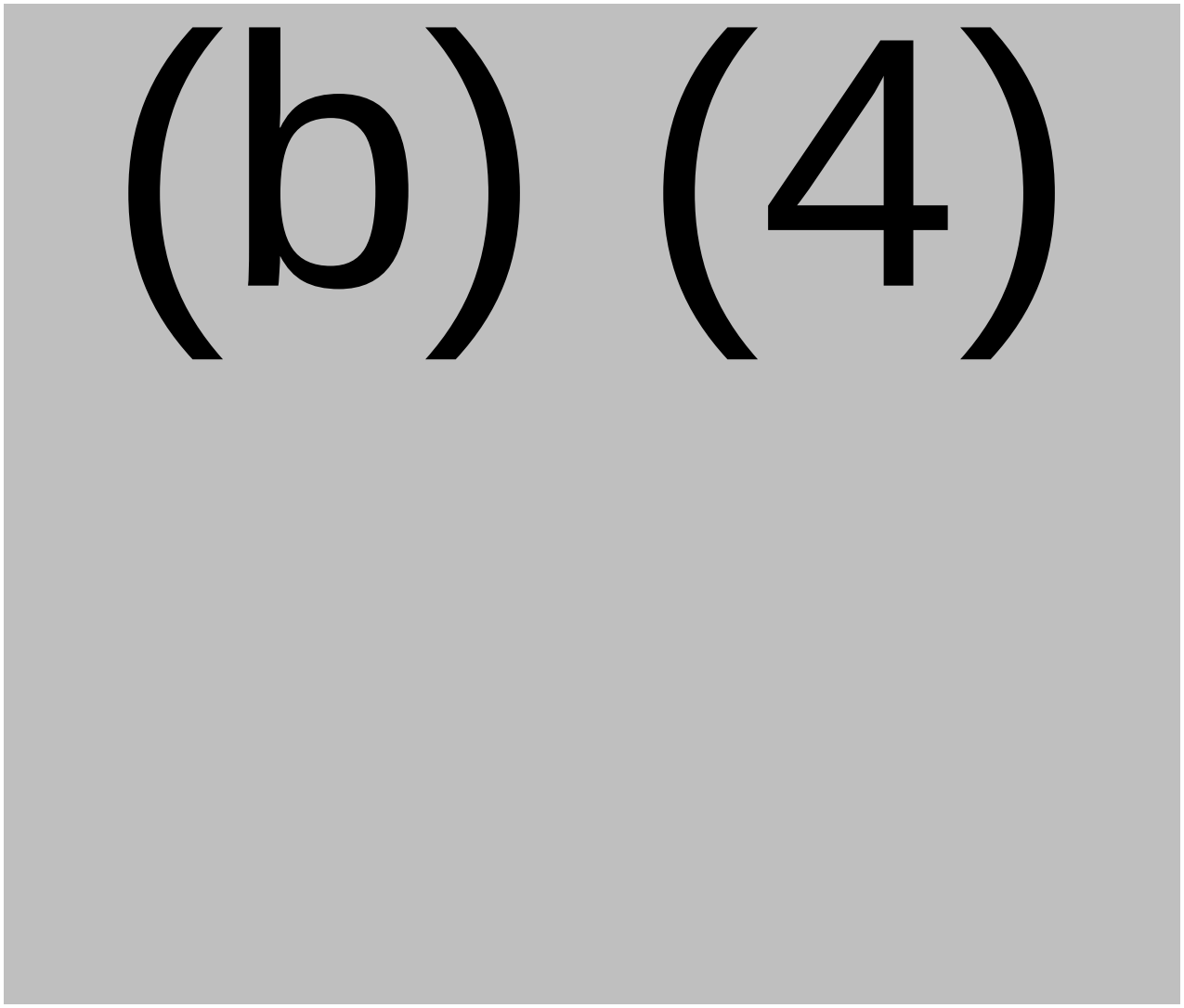
lead to systematic overestimation or underestimation of the treatment effect, of which the direction and magnitude is unknown. Further, it is difficult to interpret the magnitude of the effect using PRO data since this was a parallel arm trial with no comparator (i.e., standard of care or placebo) due to ethical considerations. Therefore, it is challenging to assess whether the treatment had an effect on a PRO endpoint(s).

(b) (4)

- ***Patients were asked to report on their pain medication use within the past 2-weeks at baseline, weeks 4, 26, 39, and 52. While the proportion of those who used 'none' increased from baseline (n=97/133, 72.9%) to week 52 (n=105/131, 80.2%), there is still roughly 20% who continued to use pain medicine despite receiving treatment and this may influence responses on a PRO evaluating pain.***
- ***The PROMIS Pain Intensity 3a Item (worst pain) is very weakly correlated to ABR (r=0.17) indicating that change in pain intensity is disassociated from other study measures. For example, a large proportion of patients (>20%) had low annualized bleeding rates (ABR) and moderate to low pain (Scores of 1-3) at baseline, while a large proportion of patients (>20%) indicated a low pain score of 1 with a 0 ABR at 52 weeks.***

(b) (4)

1 Module 2.7.3, Summary of Clinical Efficacy, Section 1.2.3.2.1.5 Quality of Life
2 Module 2.7.3, Summary of Clinical Efficacy, Section 2.2.5.1.3 Quality of life assessments



6 CLINICAL OUTCOME ASSESSMENT(S)

6.1 Clinical Outcome Assessment Description(s)

Haemophilia Quality of Life Questionnaire for Adults (Haem-A-QoL) – Physical Health Subscale

The Haem-A-QoL questionnaire was developed for participants (≥17 years of age) to evaluate physical health impairments attributed to hemophilia. The Physical Health domain includes 5 items (painful swellings, pain in joints, pain on movement, and difficulty walking, and needed more time to get ready because of the condition). Items are rated using a verbal rating scale that includes: 'Never,' 'Rarely,' 'Sometimes,' 'Often,' and 'All the

3 Module 2.7.3, Summary of Clinical Efficacy, Section 2.2.3, Table 12 - Results of the primary and secondary efficacy endpoints analyzed as part of the hierarchical testing procedure demonstrating efficacy of BIVV001.

time.' The recall period is "in the past 4 weeks." See Section 7.1.1 for a copy of the Haem-A-QoL Physical Health (PH) Subscale.

Patient-Reported Outcomes Measurement Information System Short Form v 1.0 (PROMIS-SF v 1.0) Pain Intensity 3a

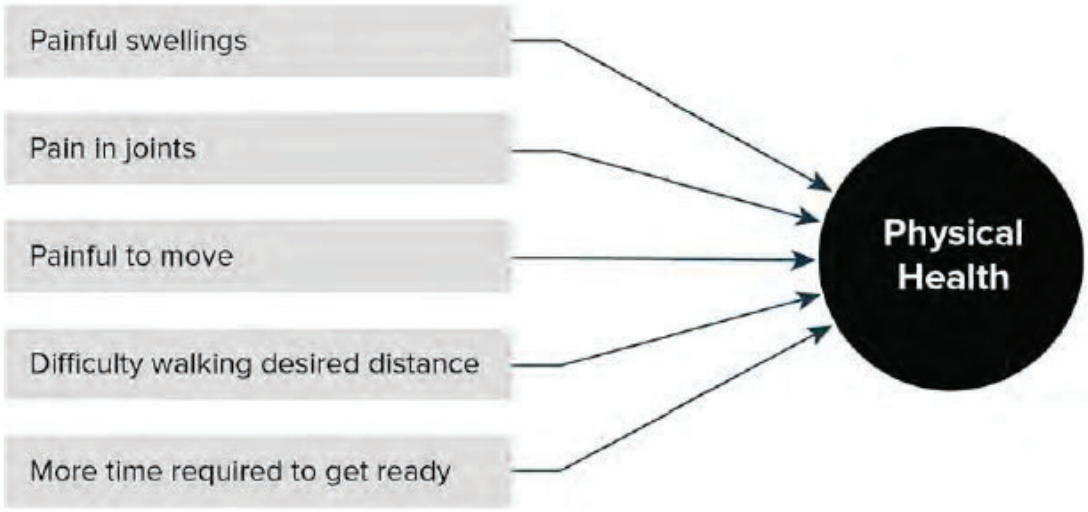
The PROMIS Pain Intensity 3a is a patient-completed questionnaire developed to evaluate pain intensity using 3 items on a 5-point Likert scale that includes the following response options: "Had no pain," "Mild," "Moderate," "Severe," and "Very severe." The first 2 items of the PROMIS Pain Intensity 3a assess peak and average pain intensity over the past 7 days. The last item asks patients to rate their pain intensity "right now." See Section 7.1.2 for a copy of the PROMIS Pain Intensity 3a measure.

Reviewer's comment(s):
All XTEND-1 patient reported outcome (PRO) measures were administered electronically using an electronic tablet. For scoring information, see Section 6.3.

6.2 Conceptual Framework(s)

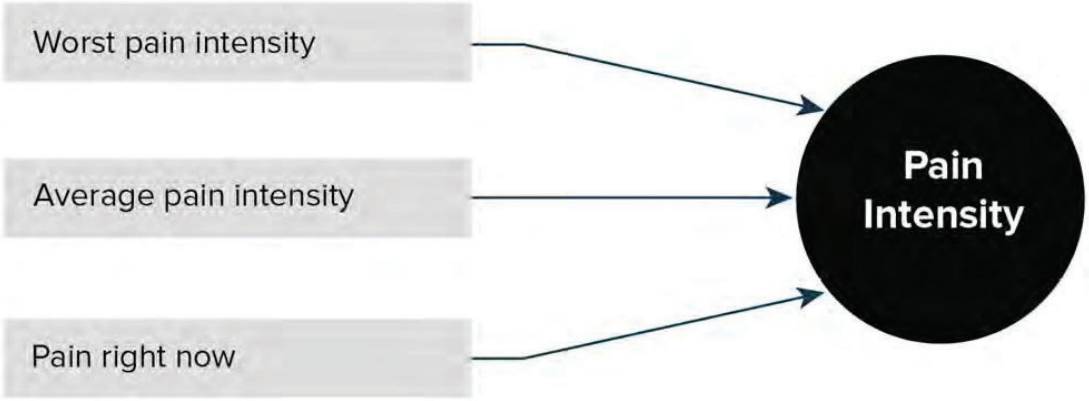
The conceptual frameworks for the Haem-A-QoL PH and PROMIS Pain Intensity 3a measures are shown in Figures 2 and 3, respectively.

Figure 2. Haem-A-QoL Physical Health (PH) Subscale (Recall period: past 4-weeks)



Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.4: Conceptual Framework: Haem-A-QoL PH, Figure 3 (p. 23 of 2016), dated 9 June 2022.

Figure 3. PROMIS Pain Intensity 3a (Recall period: past 7-days)



PROMIS = Patient-Reported Outcomes Measurement Information System.
Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 4.4: Conceptual Framework: PROMIS Pain Intensity 3a, Figure 9 (p. 63 of 2016), dated 9 June 2022.

Reviewer's comment(s):
The Haem-A-QoL PH conceptual framework subscale (Figure 2) appears to include concepts that pertain to impacts or interferences rather than symptoms. For instance, 'Walking desired distance' may be perceived as an impact and 'More time required to get ready' may be perceived as an interference.

6.3 Scoring Algorithm

Haem-A-QoL PH subscale

The Haem-A-QoL PH subscale includes five items that uses a 5-point verbal rating scale ranging from "never" (1) to "all of the time" (5). Subscale raw scores are the sums of the corresponding item responses; a score is calculated when at least 50% of the items corresponding to that subscale (or total) are non-missing. The raw score is then transformed to a scale of 0 to 100. The transformed scores for the Haem-A-QoL PH, total and other subscales range from 0 (best quality of life) to 100 (worst quality of life). Higher scores indicate lower quality of life. Transformed scale scores are computed using the following formula:

$$\frac{100 \times (raw\ score - minimal\ possible\ score)}{possible\ range\ of\ scores}$$

PROMIS Pain Intensity 3a

The PROMIS Pain Intensity 3a consists of 3 items with each item rated on a 5-point verbal rating scale ranging from 1 to 5. Summation of the three item scores yields the total score, which ranges from 3 to 15 points; higher scores indicate greater levels of pain intensity. The raw total score (ranging from 3 to 15) can be converted into a T-score (mean = 50, SD = 10 in a normative sample representative of the US general population) using the version 1 score conversion table obtained from HealthMeasures Scoring Service. In the XTEND-1 study, responses to all 3 items will be required for a non-missing T-score. The PROMIS Pain Intensity 3a Item 1 (Worst Pain, ranging from 1-5 points) and T-scores were used to compute endpoints for the XTEND-1 study.

6.4 Content Validity

The applicant completed the following activities to provide evidence to evaluate content validity of the Haem-A-QoL PH subscale PROMIS Scale v1.0 Pain Intensity 3a Item (Worst Pain):

- Review of the Literature
- Optional Exit Interviews with XTEND_1 Clinical Trial participants

Summary of findings:

Review of Literature

- Review of the Voice of the Patient Report as a part of the Patient Focused Drug Development (PFDD) initiative on heritable bleeding disorders (including hemophilia) held in September 2014[4].
- Review of published literature from the development of the Haem-A-QoL and its content validity[5-6].
- Conducted confirmatory literature review from other programs assessing performance of the Haem-A-QoL Physical Health subscale in patients with Hemophilia A[7].
- Review of published literature from the development of the PROMIS Pain 3a scale[8,9].
- Review of the literature assessing prevalence of pain and its relevance and importance to individuals with hemophilia[10,11,12].

Optional Exit Interviews

- Twenty-seven adult patients (age range:19 -73 years) from the XTEND-1 study participated in semi-structured interviews from 6 countries (Argentina, France, Italy, South Korea, UK, and US) within 6-months of exiting study but before end of study. Seventeen (n=17/29; 58.6%) exit interview participants were from Arm A while the remaining were from Arm B (n=12/29; 41.4%)
- Twenty-seven (n=27/27, 100%) of participants reported experiencing some level of hemophilia-related pain (i.e., joint pain) before starting the XTEND-1 study.

---

[4] FDA. Food and Drug Administration. The voice of the patient: a series of reports from the U.S. Food and Drug Administration's (FDA's) patient-focused drug development initiative hemophilia A, hemophilia B, von Willebrand disease and other heritable bleeding disorders. 2016. https://www.fda.gov/files/about%20fda/published/The-Voice-of-the- Patient--Hemophilia-A--Hemophilia-B--von-Willebrand-Disease-and-Other-Heritable- Bleeding-Disorders.pdf.

[5] von Mackensen S, Gringeri A. Quality of life in hemophilia. In: Preedy VR, Watson RR, editors. Handbook of Disease Burdens and Quality of Life Measures. New York, NY: Springer; 2010.

[6] von Mackensen S, Gringeri A, Haem-A-QoL study Group. Health-related quality of life in adult patients with haemophilia – Assessment with a new disease-specific questionnaire (Haem-A-QoL). J Thrombo Haem. 2005a;3(Suppl 1):P0813.

[7] Skinner MW, Negrier C, Paz-Priel I, Chebon S, Jimenez-Yuste V, Callaghan MU, et al. The effect of emicizumab prophylaxis on long-term, self-reported physical health in persons with haemophilia A without factor VIII inhibitors in the HAVEN 3 and HAVEN 4 studies. Haemophilia. 2021 Sep;27(5):854-65. doi:http://dx.doi.org/10.1111/hae.14363.

[8] Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al. PROMIS((R)) Adult Health Profiles: efficient short-form measures of seven health domains. Value Health. 2019 May;22(5):537-44. doi:http://dx.doi.org/10.1016/j.jval.2019.02.004.

[9] Chen W-H, Revicki D, Amtmann D, Jensen MP, Keefe FJ, Cella D. University of Washington: Health Aging and Physical Disability: Rehabilitation Research and Training Center. Development and Analysis of PROMIS Pain Intensity Scale. 2012. https://agerrtc.washington.edu/index.php?q=node/55. Accessed 31 January 2022

[10] Paredes AC, Teixeira P, Almeida A, Pinto PR. Prevalence and interference of chronic pain among people with Hemophilia: a systematic review and meta-analysis. J Pain. 2021 Oct;22(10):1134-45. doi:http://dx.doi.org/10 1016/j.jpain.2021.03.157.

[11] Witkop M, Neff A, Buckner TW, Wang M, Batt K, Kessler CM, et al. Self-reported prevalence, description and management of pain in adults with haemophilia: methods, demographics and results from the Pain, Functional Impairment, and Quality of life (P-FiQ) study. Haemophilia. 2017 Jul;23(4):556-65. doi:http://dx.doi.org/10.1111/hae.13214.

[12] Taylor S, Toye F, Donovan-Hall M, Barker K. Past the tipping point: a qualitative study of the views and experiences of men with haemophilia regarding mobility, balance, and falls. Disabil Rehabil. 2021 Oct 15:1-9. doi:http://dx.doi.org/10.1080/09638288.2021.1988731.

- A majority of participants (n=24/29, 82.8%) reported that they were 'very satisfied' with the treatment and all (n=29/29, 100%) indicated they preferred the experimental treatment over their previous hemophilia treatment.
- Below are the results from the optional exit interviews for each PRO measure:
  a) Haem-A-QoL
     - Most participants (n=24/27; 88.9%) reported improvement in at least 1 of the Haem-A-QoL PH concepts during the XTEND-1 trial.
     - More than 80% of participants (n≥22/27) reported improvements in joint pain, the ability to move without pain, and painful swelling compared to the start of the trial.
     - Improvements were reported by a large portion of exit interview participants for each of the Haem-A-QoL PH items: 'painful swelling' (n-20/24; 83.3%), 'pain in joints' (n=24/27; 88.9%), 'ability to move without pain (n=22/25; 88.0%), 'ability to walk desired distance' (n=18/24, 75.0%), 'time to get ready' (n=9/12, 75.0%), respectively.
  b) PROMIS Scale v1.0 Pain Intensity 3a Item
     - All participants (n=27/27; 100%) indicated questions for the PROMIS Pain Intensity 3a were relevant, clear, and easy to answer with the provided answer options; and recall period was appropriate (7-days).
     - While all participants (n=27/27; 100%) demonstrated ability to select responses for all 3-items, a few (n=3/27; 11.1%) found the items addressing 'worst pain' and 'current pain' (recall period 'right now') easier to answer than the item asking participants to average their pain over 7-days.
     - All participants (n=27/27; 100%) stated that any positive change (i.e., 1 point or more) would be meaningful.
     - A majority of interview participants (n=24/27; 88.9%) noted an improvement in pain during the XTEND-1 study. The remaining participants (n=3/27; 11.1%) noted no changes in pain intensity attributing no improvements to cumulative damage sustained over the years from repeated bleeds.
     - Most patients reported that pain was closely related to additional symptoms and functional impacts, including inability to move without pain (n=25/27; 92.6%), painful [joint] swelling (n=24/27; 88.9%); inability to walk without pain and/or walk desired distances (n =24/27; 88.9%).

Reviewer's comment(s):

- ***Applicant did not provide copies of exit interview transcripts. Transcripts were not requested because there were overarching concerns with the study design (i.e., open label) that took precedent when interpreting the results of the XTEND-1 secondary endpoint data.***
- ***Optional exit interviews were used to evaluate both content validity and clinically meaningful within patient change for both the Haem-A-QoL PH and PROMIS Pain Intensity 3a item (Worst Pain). Further details on clinically meaningful within patient change based on anchor-based analyses can be found in section 6.6.1.***

6.5 Other Measurement Properties

Measurement properties of both the Haem-A-QoL and PROMIS Pain Intensity 3a item 1 (Worst Pain) consisted of a review of literature and analysis of pooled data from Arm A and Arm B of the XTEND-1 study. However, only data from patients aged ≥ 17 years were used in the psychometric analyses of the PROMIS Pain Intensity 3a item 1 (Worst Pain).

Approximately 147 patients (aged ≥ 12 years) pooled across Arm A and Arm B from the XTEND-1 study who received at least 1 dose of the investigational treatment were used to evaluate the psychometric properties (construct validity, test-retest reliability, etc.). The majority of patients were male (n=145, 99.3%), white (n=91, 61.9%) and neither Hispanic nor Latino (n=114, 78.1%). The top three most reported affected target joints were: right ankle (n=29, 64.4%), left ankle (n=24, 53.3%) and right knee (n=21, 46.7%), respectively.

Reviewer's comment(s):

- ***The PROMIS Pain Intensity 3a measurement properties were analyzed using pooled data from Arm and B of the XTEND-1 study and included all three items rather than the single item (Worst Pain).*** (b) (4)
- ***For the purposes of this review, other measurement properties were evaluated at the prespecified timepoints (i.e., baseline, Weeks 26 and/or 52) for the two PRO measures using the PGIS-Joint and PGIS-activity as reference measures. The other measures (i.e., EQ-5D-5L) were not used, as the content validity was not fully evaluated in this study population.***

Descriptive Statistics

Item-level response frequencies were analyzed to evaluate floor/ceiling effects for the Haem-A-QoL PH using pooled sample from Arms A and B from the XTEND-1 study from change at baseline to week 52. According to the Applicant, "no ceiling and floor effects were observed across the Haem-A-QoL PH 5-items when the percentages of the best (score 1, "Never") or worst (score 5, "All the time") responses were calculated using the Applicant's pre-selected 40% criterion (Table 2).

The Agency requested additional analyses be performed assessing item-level distributions of the Haem-A-QoL PH for Arm A from the XTEND-1 study (Information Request; SDN 20, received date 02 December 2022). Refer to reviewer's comment below.

Table 2. Haem-A-QoL PH Item-Level Descriptive Statistics: Pooled Arm A and B (N=126), XTEND-1 Study

	Baseline distribution				CFB to Week 52		
Haem-A-QoL PH	n[a]	Mean (SD), median	Score min %, max %	Missing (%)	n[a]	Mean (SD), median	Missing (%)
Item 1. Swelling	126	2.4 (1.19), 2.0	30.2, 4.8	0 (0.0)	120	−0.6 (1.23), 0.0	6 (4.8)
Item 2. Pain in Joints	126	2.9 (1.20), 3.0	18.3, 8.7	0 (0.0)	120	−0.3 (1.03), 0.0	6 (4.8)
Item 3. Painful to Move	126	2.7 (1.13), 3.0	19.0, 5.6	0 (0.0)	120	−0.3 (0.99), 0.0	6 (4.8)
Item 4. Difficulty Walking	126	2.6 (1.26), 3.0	25.4, 10.3	0 (0.0)	120	−0.3 (1.25), 0.0	6 (4.8)
Item 5. More Time to Get Ready	126	2.2 (1.12), 2.0	34.1, 3.2	0 (0.0)	120	−0.4 (1.02), 0.0	6 (4.8)

CFB = change from Baseline; max = maximum; min = minimum; Haem-A-QoL = Haemophilia Quality of Life; PH = physical health; SD = standard deviation.
Note: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.
[a] Sample sizes exclude participants with missing data.
Source: Applicant's PRO Dossier for the Haem-A-QoL, Section 3.6.2: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 8. (pg. 38 of 2016), dated 9 June 2022.

Reviewer's comment(s):
As shown in Table 3, there are concerns with potential floor effects when using the Arm A sample to assess efficacy on the secondary endpoint (Haem-A-QoL) (Information Request; SDN 20, received date 02 December 2022). A higher proportion of patients indicated that they rated items on the Haem-A-QoL PH as 'never,' 'rarely,' or 'sometimes' at baseline, suggesting less room for change in these scores in response to treatment. For instance, more than 50% (n=103) selected "never,' 'rarely' for items 1 and 3-5. A low baseline severity and little change over 52 weeks on the PGIS-Joint (reference measure), as depicted in Table 4, suggests that the measure may not be sensitive to change, or that there were few patients who experienced a meaningful change ('improvement' or 'worsening') during the open-label observation period.

Table 3. Distribution of items at baseline on Haem-A-QoL of Arm A patients in the XTEND-1 Study

Haem-A-QoL - Physical Health items, n(%)	Arm A: Prophylaxis (N=103)
Item 1 - Swelling	
Baseline	
Never	36/103 (35.0)
Rarely	23/103 (22.3)
Sometimes	29/103 (28.2)
Often	13/103 (12.6)
All the time	2/103 (1.9)
Missing	0/103
Item 2 - Pain in Joints	
Baseline	
Never	20/103 (19.4)
Rarely	14/103 (13.6)
Sometimes	36/103 (35.0)
Often	25/103 (24.3)
All the time	8/103 (7.8)
Missing	0/103
Item 3 - Painful to Move	
Baseline	
Never	20/103 (19.4)
Rarely	26/103 (25.2)
Sometimes	34/103 (33.0)
Often	17/103 (16.5)
All the time	6/103 (5.8)
Missing	0/103
Item 4 - Difficulty Walking	
Baseline	
Never	28/103 (27.2)
Rarely	25/103 (24.3)
Sometimes	28/103 (27.2)
Often	11/103 (10.7)
All the time	11/103 (10.7)
Missing	0/103
Item 5 - More Time to Get Ready	
Baseline	
Never	40/103 (38.8)
Rarely	22/103 (21.4)
Sometimes	28/103 (27.2)
Often	10/103 (9.7)
All the time	3/103 (2.9)
Missing	0/103

The Haem-A-QoL psychometric analysis sample was defined as all participants >=17 years old who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data, and had at least 1 evaluable post-Day 1 Haem-A-QoL data during the treatment period.
Source: Response to FDA Information Request, Table 1. (pg. 9-10 of 42), dated 16 November 2022.

Table 4. Change from Baseline to Week 52 in PGIS Joint according to PGIS Joint at baseline-Arm A, XTEND-1 Study

	Change from baseline in PGIS Joint at week 52								
PGIS Joint at baseline	Improved 4 categories (N=0)	Improved 3 categories (N=2)	Improved 2 categories (N=8)	Improved 1 category (N=23)	No Change (N=66)	Worsened 1 category (N=14)	Worsened 2 categories (N=1)	Worsened 3 categories (N=1)	Worsened 4 categories (N=0)
None	0/0	0/2	0/8	0/23	31/66 (47.0)	6/14 (42.9)	0/1	1/1 (100)	0/0
Mild	0/0	0/2	0/8	8/23 (34.8)	16/66 (24.2)	6/14 (42.9)	1/1 (100)	0/1	0/0
Moderate	0/0	0/2	4/8 (50.0)	12/23 (52.2)	13/66 (19.7)	2/14 (14.3)	0/1	0/1	0/0
Severe	0/0	1/2 (50.0)	3/8 (37.5)	2/23 (8.7)	5/66 (7.6)	0/14	0/1	0/1	0/0
Very severe	0/0	1/2 (50.0)	1/8 (12.5)	1/23 (4.3)	1/66 (1.5)	0/14	0/1	0/1	0/0

% are calculated using the number of patients in each change category as denominator

Source: Response to FDA Information Request, Table 10. (pg. 38 of 42), dated 16 November 2022.

Test-Retest Reliability

Test-retest reliability was analyzed by the Applicant using Haem-Q-QoL scores from a stable subgroup of patients, defined as patients with no change in PGIS items ('Joint' and 'Physical Activity'), from Week 26 to Week 52. Intraclass correlation coefficients (ICCs) were computed using a two-way mixed-effects ANOVA with absolute agreement for single measures and criterion ≥ 0.70 (Nunnally and Bernstein, 1994) for multi-item scales, indicating acceptable test-retest reliability per the Applicant as shown in Table 5 for the Haem-A-QoL PH.

As shown in Table 5 and Table 6 by the Applicant, ICCs were higher (ICC > 0.70) for Haem-A-QoL PH and lower (ICC < 0.70) for the PROMIS single item using the PGIS Joint, respectively.

Table 5. Test-Retest Reliability Results of Haem-A-QoL PH Score of pooled Arm A and B of XTEND-1 Study

	Test-retest Reliability[a] (95% CI), n	
Haem-A-QoL	Stable PGIS-Joint	Stable PGIS-Activity
PH	0.85 (0.77-0.90), 80	0.78 (0.68-0.86), 72

CI = confidence interval; Haem-A-QoL = Haemophilia Quality of Life; ICC = intraclass correlation coefficient; PGIS = Patient Global Impression of Severity; PH = physical health.

Note: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.

[a] ICCs were computed for all scores.

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.6.2: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 11. (pg. 41 of 2016), dated 9 June 2022.

Table 6. Test-Retest Reliability of PROMIS Pain Intensity 3a Item (Worst pain) and T-score of pooled Arm A and B of XTEND-1 Study

	Test-retest Reliability[a] (95% CI), n	
Sample/PROMIS Pain Intensity 3a	**PGIS-joint stable from week 26 to week 52 subset**	**PGIS-activity stable from week 26 to week 52 subset**
Pooled Arms A and B (N = 147)		
Total (raw)	0.70 (0.59-0.79), 96	0.65 (0.51-0.76), 87
Total (T-score)	0.69 (0.57-0.78), 96	0.65 (0.51-0.76), 87
Item 1. Worst Pain	0.57 (0.39-0.76), 96	0.57 (0.40-0.75), 87
Item 2. Average Pain	0.58 (0.39-0.77), 96	0.50 (0.32-0.67), 87
Item 3. Current Pain	0.68 (0.56-0.80), 96	0.70 (0.56-0.84), 87
Arm A (N = 124)		
Total (raw)	0.73 (0.60-0.81), 81	0.77 (0.65-0.85), 73
Total (T-score)	0.71 (0.59-0.81), 81	0.75 (0.62-0.83), 73
Item 1. Worst Pain	0.61 (0.41-0.80), 81	0.69 (0.53-0.86), 73
Item 2. Average Pain	0.62 (0.42-0.82), 81	0.61 (0.41-0.81), 73
Item 3. Current Pain	0.70 (0.57-0.82), 81	0.76 (0.64-0.88), 73

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.6.2: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table B-8a, (pg. 1087of 2016), dated 9 June 2022.

Construct Validity

A Pearson's correlation was calculated by the Applicant to evaluate the associations between Haem-A-QoL PH subscale and total scores and scores on other measures at baseline and week 52. The same calculation was done for the PROMIS Pain Intensity 3a Worst pain Item and total score. The results from the Applicant are reported in Table 7 and Table 8, respectively.

Reviewer's comment(s):
Moderate correlations (i.e., 0.3 to 0.7) were seen for both the Haem-A-QoL PH and the PROMIS Pain Intensity 3a (Worst Pain) item and the reference measures.

Table 7. Summary of Convergent Validity Results of Haem-A-QoL PH for pooled Arm A and B from XTEND-1 Study.

	Correlation between Haem-A-QoL PH with supporting measures[a], (n)	
Supporting measure	**Baseline**	**Week 52**
PROMIS-SF Pain Interference 6a T-score	0.71* (125)	0.82* (119)
PROMIS-SF Physical Function 6b T-score	−0.73* (125)	−0.70* (119)
PGIS-Joint	0.72* (123)	0.77* (119)
PGIS-Activity	0.68* (123)	0.73* (119)
HAL Total	−0.78* (125)	−0.69* (118)
EQ-VAS	−0.50* (123)	−0.55* (119)
EQ-5D-5L Mobility	0.76* (123)	0.67* (119)
EQ-5D-5L Usual Activities	0.78* (123)	0.66* (119)
EQ-5D-5L Pain/Discomfort	0.69* (123)	0.72* (119)
HJHS Total	0.55* (110)	0.47* (104)

* $P < 0.05$ for H_0: $\rho = 0$.

Haem-A-QoL = Haemophilia Quality of Life; HAL = Haemophilia Activities List; HJHS = Hemophilia Joint Health Score; PGIS = Patient Global Impression of Severity; PH = physical health; PROMIS = Patient-Reported Outcomes Measurement Information System; SF = Short Form; VAS = visual analog scale.

Note: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.

[a] Polyserial correlations were computed with the ordinal scores of supportive measures, and Pearson correlations were computed with the continuous scores of supportive measures.

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.6.2: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 12. (pg. 42 of 2016), dated 9 June 2022.

Table 8. Convergent Validity Results of Key PROMIS Pain Intensity 3a Scores: Pooled Arm A and B (N=147), XTEND-1 Study

	Correlation With PROMIS Pain Intensity 3a scores[a], (n)			
	Worst Pain		T-score	
Supporting measure	Baseline	Week 52	Baseline	Week 52
PROMIS-SF Pain Interference 6a T-score	0.60* (125)	0.73* (119)	0.68* (125)	0.74* (119)
PROMIS-SF Physical Function 6b T-score	−0.42* (125)	−0.52* (119)	−0.47* (125)	−0.55* (119)
PGIS-Joint	0.72* (143)	0.72* (140)	0.71* (143)	0.71* (140)
PGIS-Activity	0.57* (143)	0.64* (140)	0.59* (143)	0.66* (140)
HAL Total	−0.44* (125)	−0.53* (118)	−0.51* (125)	−0.55* (118)
EQ-VAS	−0.46* (143)	−0.51* (140)	−0.50* (143)	−0.55* (140)
EQ-5D-5L Mobility	0.53* (143)	0.52* (140)	0.54* (143)	0.54* (140)
EQ-5D-5L Usual Activities	0.49* (143)	0.53* (140)	0.51* (143)	0.56* (140)
EQ-5D-5L Pain/Discomfort	0.68* (143)	0.75* (140)	0.75* (143)	0.78* (140)
HJHS Total	0.26* (130)	0.25* (123)	0.28* (130)	0.30* (123)

*P < 0.05 for H0: ρ 0.
HAL = Haemophilia Activities List; HJHS = Hemophilia Joint Health Score; PGIS = Patient Global Impression of
Severity; PROMIS = Patient-Reported Outcomes Measurement Information System; SF = Short Form; VAS = visual analog scale.
Note: The PROMIS Pain Intensity 3a psychometric analysis sample was defined as all participants who received at
least 1 dose of the study intervention, had an evaluable Day 1 PROMIS Pain Intensity 3a data point, and had at least 1 evaluable post-Day 1 PROMIS Pain Intensity 3a data point during the treatment period.
[a] For the PROMIS Pain Intensity 3a total score, polyserial correlations are computed with the ordinal scores of supportive measures, and Pearson correlations are computed with the continuous scores of supportive measures.
For the PROMIS Pain Intensity 3a items, polychoric correlations are computed with the ordinal scores of supportive measures, and polyserial correlations are computed with the continuous scores of supportive measures.
Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 4.6: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 24. (pg. 79 of 2016), dated 9 June 2022.

Known-Groups Validity

Known-groups validity was analyzed using mean Haem-A-QoL PH subscale scores by known groups based on the PGIS ('Joint' and 'Activity'), as reference measures, at baseline and week 52 by the Applicant. As shown in Table 9 below, high mean Haem-A-QoL PH scores correspond to greater severity on the other measures at both baseline and week 52.

As shown in Table 10 for the PROMIS pain measure by the Applicant, the pattern of mean scores across the known groups were in the anticipated direction: higher (worse) PROMIS Pain Intensity 3a scores for worse PGIS-Joint rating, worse PGIS-Activity ratings.

Table 9. Summary of Known-Groups Analysis for the Haem-A-QoL PH Subscale Score at Baseline and Week 52

	Haem-A-QoL PH ANOVA results Mean (SD), median, n	
Known group	**Baseline**	**Week 52**
PGIS-Joint		
1 = No joint symptoms	13.2 (16.89), 10.0, 25	6.9 (12.89), 0.0, 36
2 = Mild joint symptoms	31.2 (14.41), 30.0, 34	27.8 (15.07), 25.0, 47
3 = Moderate joint symptoms	47.6 (19.40), 45.0, 46	44.6 (16.92), 45.0, 23
4 = Severe joint symptoms	61.5 (15.60), 60.0, 13	62.7 (18.35), 65.0, 11
5 = Very severe joint symptoms	68.0 (19.24), 70.0, 5	50.0 (0.00), 50.0, 2
ANOVA F-statistic, *P* value	27.7, $P < 0.0001$	40.1, $P < 0.0001$
PGIS-Activity		
1 = No restriction	15.2 (16.27), 10.0, 30	10.8 (14.59), 5.0, 44
2 = Mild restriction	35.9 (15.46), 35.0, 35	29.5 (15.03), 30.0, 41
3 = Moderate restriction	48.5 (21.37), 50.0, 40	44.0 (21.79), 45.0, 25
4 = Severe restriction	54.3 (15.42), 57.5, 14	66.9 (10.33), 67.5, 8
5 = Very severe restriction	77.5 (18.93), 85.0, 4	50.0 (NA), 50.0, 1
ANOVA F-statistic, *P* value	23.5, $P < 0.0001$	30.1, $P < 0.0001$
EQ-5D-5L Mobility		
1 = No problems walking	19.9 (17.40), 20.0, 49	16.3 (20.08), 10.0, 60
2 = Slight problems walking	44.7 (15.93), 45.0, 45	32.6 (15.64), 35.0, 35
3 = Moderate problems walking	53.3 (16.49), 55.0, 20	46.5 (15.98), 45.0, 17
4 = Severe problems walking	74.4 (17.40), 80.0, 9	65.7 (16.18), 70.0, 7
5 = Unable to walk	NA, 0	NA, 0
ANOVA F-statistic, *P* value	41.3, $P < 0.0001$	25.0, $P < 0.0001$
EQ-5D-5L Usual Activities		
1 = No problems doing usual activities	25.7 (19.46), 25.0, 68	20.0 (19.94), 15.0, 82
2 = Slight problems doing usual activities	46.9 (15.24), 45.0, 39	41.5 (16.23), 40.0, 26
3 = Moderate problems doing usual activities	67.7 (12.18), 65.0, 13	60.0 (17.95), 57.5, 10
4 = Severe problems doing usual activities	86.7 (5.77), 90.0, 3	50.0 (NA), 50.0, 1
5 = Unable to do usual activities	NA, 0	NA, 0
ANOVA F-statistic, *P* value	35.1, $P < 0.0001$	19.1, $P < 0.0001$

EQ-5D-5L Pain/Discomfort		
1 = No pain or discomfort	20.0 (18.00), 20.0, 45	12.7 (18.32), 5.0, 47
2 = Slight pain or discomfort	43.7 (17.61), 45.0, 50	31.6 (15.63), 35.0, 54
3 = Moderate pain or discomfort	55.9 (21.08), 57.5, 22	58.6 (17.03), 60.0, 14
4 = Severe pain or discomfort	65.0 (11.18), 65.0, 5	62.5 (15.00), 60.0, 4
5 = Extreme pain or discomfort	80.0 (NA), 80.0, 1	NA, 0
ANOVA F-statistic, *P* value	21.5, $P < 0.0001$	34.5, $P < 0.0001$

ANOVA = analysis of variance; Haem-A-QoL = Haemophilia Quality of Life; NA = not applicable; PGIS = Patient Global Impression of Severity; PH = physical health; SD = standard deviation.

Note: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.6.2: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 13. (pg. 43-44 of 2016), dated 9 June 2022.

Table 10. Known-Groups Validity Results of Key PROMIS Pain Intensity 3a Scores: Pooled Arm A and Arm B (N =147), XTEND-1 Study

	PROMIS Pain Intensity 3a ANOVA results Mean (SD), median, n			
	Worst Pain		T-score	
Known group	**Baseline**	**Week 52**	**Baseline**	**Week 52**
PGIS-Joint Known Group				
No joint symptoms	1.5 (0.81), 1.0, 40	1.4 (0.81), 1.0, 54	34.6 (5.52), 30.7, 40	33.8 (5.55), 30.7, 54
Mild joint symptoms	2.3 (0.85), 2.0, 38	2.1 (0.87), 2.0, 48	43.0 (6.01), 43.5, 38	40.4 (6.23), 40.2, 48
Moderate joint symptoms	3.2 (0.82), 3.0, 47	3.2 (0.91), 3.0, 25	48.2 (6.10), 49.4, 47	48.7 (6.50), 49.4, 25
Severe joint symptoms	3.5 (0.88), 4.0, 13	3.7 (1.10), 4.0, 11	49.7 (5.85), 49.4, 13	51.6 (9.01), 54.5, 11
Very severe joint symptoms	3.4 (1.34), 4.0, 5	3.0 (1.41), 3.0, 2	52.7 (12.39), 57.5, 5	49.0 (7.78), 49.0, 2
ANOVA F-statistic, *P* value	29.5, $P < 0.0001$	27.3, $P < 0.0001$	33.8, $P < 0.0001$	35.0, $P < 0.0001$
PGIS-Activity Known Group				
No restriction	1.7 (0.99), 1.0, 45	1.5 (0.86), 1.0, 62	36.9 (7.03), 36.3, 45	34.9 (6.24), 30.7, 62
Mild restriction	2.5 (0.98), 2.5, 38	2.3 (0.94), 2.0, 44	43.2 (6.57), 43.5, 38	41.6 (6.08), 43.5, 44
Moderate restriction	3.1 (0.95), 3.0, 42	2.9 (1.20), 3.0, 25	47.9 (7.13), 49.4, 42	46.6 (9.25), 49.4, 25
Severe restriction	3.0 (1.18), 3.5, 14	4.1 (0.64), 4.0, 8	46.5 (8.25), 47.8, 14	55.5 (5.14), 54.5, 8
Very severe restriction	4.0 (0.00), 4.0, 4	2.0 (NA), 2.0, 1	57.5 (2.45), 57.5, 4	43.5 (NA), 43.5, 1
ANOVA F-statistic, *P* value	13.6, $P < 0.0001$	19.1, $P < 0.0001$	18.7, $P < 0.0001$	25.5, $P < 0.0001$
EQ-5D-5L Mobility Known Group				
No problems walking	2.0 (1.01), 2.0, 69	1.8 (1.07), 1.0, 80	39.2 (7.56), 40.2, 69	36.8 (7.86), 30.7, 80
Slight problems walking	2.9 (0.99), 3.0, 45	2.5 (0.84), 2.0, 36	46.0 (6.90), 46.3, 45	43.4 (5.90), 43.5, 36
Moderate problems walking	3.0 (1.05), 3.0, 20	2.9 (1.32), 3.0, 17	47.8 (7.96), 49.4, 20	46.8 (10.04), 49.4, 17
Severe problems walking	3.6 (1.24), 4.0, 9	3.4 (1.40), 4.0, 7	51.7 (9.20), 54.5, 9	49.5 (9.26), 54.5, 7
Unable to walk	NA, 0	NA, 0	NA, 0	NA, 0
ANOVA F-statistic, *P* value	13.0, $P < 0.0001$	9.9, $P < 0.0001$	15.1, $P < 0.0001$	14.7, $P < 0.0001$

EQ-5D-5L Usual Activities Known Group				
No problems doing usual activities	2.2 (1.05), 2.0, 88	1.9 (1.02), 2.0, 103	40.7 (7.96), 43.5, 88	37.8 (7.45), 36.3, 103
Slight problems doing usual activities	2.8 (1.14), 3.0, 39	2.8 (1.18), 3.0, 26	45.8 (7.91), 46.3, 39	46.1 (8.91), 46.3, 26
Moderate problems doing usual activities	3.4 (0.96), 4.0, 13	3.6 (1.07), 4.0, 10	50.5 (6.87), 52.1, 13	51.2 (7.61), 54.5, 10
Severe problems doing usual activities	4.0 (0.00), 4.0, 3	2.0 (NA), 2.0, 1	55.5 (1.73), 54.5, 3	43.5 (NA), 43.5, 1
Unable to do usual activities	NA, 0	NA, 0	NA, 0	NA, 0
ANOVA F-statistic, *P* value	8.9, $P < 0.0001$	11.1, $P < 0.0001$	10.7, $P < 0.0001$	15.1, $P < 0.0001$
EQ-5D-5L Pain/Discomfort Known Group				
No pain or discomfort	1.8 (0.99), 1.0, 62	1.5 (0.83), 1.0, 65	37.1 (6.81), 36.3, 62	34.2 (5.72), 30.7, 65
Slight pain or discomfort	2.8 (0.94), 3.0, 53	2.5 (0.90), 2.0, 55	46.1 (6.03), 46.3, 53	42.9 (6.13), 43.5, 55
Moderate pain or discomfort	3.4 (0.85), 3.0, 22	3.8 (0.86), 4.0, 16	50.6 (5.92), 52.1, 22	52.6 (5.15), 54.5, 16
Severe pain or discomfort	4.0 (0.00), 4.0, 5	3.8 (1.26), 4.0, 4	55.2 (2.30), 54.5, 5	55.0 (9.77), 54.5, 4
Extreme pain or discomfort	4.0 (NA), 4.0, 1	NA, 0	60.5 (NA), 60.5, 1	NA, 0
ANOVA F-statistic, *P* value	18.8, $P < 0.0001$	37.4, $P < 0.0001$	31.0 (< 0.0001)	57.3, $P < 0.0001$

ANOVA = analysis of variance; PGIS = Patient Global Impression of Severity; PROMIS = Patient-Reported Outcomes Measurement Information System; SD = standard deviation.
Note: The PROMIS Pain Intensity 3a psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 PROMIS Pain Intensity 3a data point, and had at least 1 evaluable post-Day 1 PROMIS Pain Intensity 3a data point during the treatment period.
Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 4.6.2.9: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 25. (pg. 81-82 of 2016), dated 9 June 2022.

Ability to Detect Change

Ability to detect change was supported by the direction and magnitude of correlations between change in the Haem-A-QoL PH and PROMIS Pain Intensity 3a and change in reference measures by the Applicant. As shown in Table 11, weak to moderate $|r| \geq 0.30$) correlations were shown in change transformed scores for the Haem-A-QoL compared to the change scores for the reference measures. Similar results were found by the Applicant for the PROMIS Pain Intensity 3a change scores compared to the reference measures, as depicted in Table 12.

Table 11. Responsiveness Correlations for Haem-A-QoL PH Transformed Scores (Pooled Arm A and B), XTEND-1 Study

	Correlations between CFB Haem-A-QoL PH and supportive measure[a] (n)	
Supportive measure	**Week 26**	**Week 52**
PROMIS-SF Pain Interference 6a T-score (CFB)	0.46* (123)	0.49* (119)
PROMIS-SF Physical Function 6b T-score (CFB)	−0.39* (123)	−0.62* (119)
PGIS-Joint (CFB)	0.36* (120)	0.47* (116)
PGIS-Activity (CFB)	0.24* (120)	0.46* (116)
PGIC-overall (at Weeks 26 or 52)	0.25* (123)	0.41* (119)
HAL Total (CFB)	−0.35* (122)	−0.42* (118)
EQ-VAS (CFB)	−0.43* (120)	−0.45* (116)
EQ-5D-5L Mobility (CFB)	0.37* (120)	0.37* (116)
EQ-5D-5L Usual Activities (CFB)	0.42* (120)	0.42* (116)
EQ-5D-5L Pain/Discomfort (CFB)	0.43* (120)	0.42* (116)
PGA (at Weeks 26 or 52)	0.15 (122)	0.02 (118)
HJHS Total (CFB)	0.20 (91)	0.26* (101)

* $P < 0.05$ for H_0: $\rho = 0$.

CFB = change from Baseline; Haem-A-QoL = Haemophilia Quality of Life; HAL = Haemophilia Activities List; HJHS = Hemophilia Joint Health Score; PGA = Physician Global Assessment; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; PH = physical health; PROMIS = Patient-Reported Outcomes Measurement Information System; SF = Short Form; VAS = visual analog scale.

Note: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.

[a] Polyserial correlations are computed with PGA and PGIC scores, and Pearson correlations are computed with the changes in other supportive measures.

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.6.2.10: Evidence of Other Measurement Properties Using Data from XTEND-1 Study, Table 14. (pg. 45 of 2016), dated 9 June 2022.

Table 12. Responsiveness Correlations for Key PROMIS Pain Intensity 3a Scores: Pooled Arm A and B (N =147), XTEND-1

	Correlation with key PROMIS Pain Intensity 3a change scores, (n)			
	Worst Pain		T-score	
Supporting measure	Week 26	Week 52	Week 26	Week 52
PROMIS-SF Pain Interference 6a T-score (CFB)	0.60* (123)	0.51* (119)	0.64* (123)	0.56* (119)
PROMIS-SF Physical Function 6b T-score (CFB)	−0.29* (123)	−0.35* (119)	−0.31* (123)	−0.40* (119)
PGIS-Joint (CFB)	0.43* (140)	0.43* (136)	0.46* (140)	0.40* (136)
PGIS-Activity (CFB)	0.22* (140)	0.32* (136)	0.25* (140)	0.35* (136)
PGIC-Overall (at Weeks 26 or 52)	0.20* (144)	0.22* (140)	0.24* (144)	0.24* (140)
HAL Total (CFB)	−0.23* (122)	−0.33* (118)	−0.23* (122)	−0.31* (118)
EQ-VAS (CFB)	−0.34* (140)	−0.26* (136)	−0.40* (140)	−0.34* (136)
EQ-5D-5L Mobility (CFB)	0.23* (140)	0.09 (136)	0.28* (140)	0.18* (136)
EQ-5D-5L Usual Activities (CFB)	0.27* (140)	0.19* (136)	0.27* (140)	0.27* (136)
EQ-5D-5L Pain/Discomfort (CFB)	0.34* (140)	0.30* (136)	0.45* (140)	0.40* (136)
PGA (at Weeks 26 or 52)	0.06 (142)	−0.09 (138)	0.06 (142)	−0.05 (138)
HJHS Total (CFB)	0.11 (111)	0.12 (120)	0.12 (111)	0.11 (120)

P < 0.05 for H0: ρ 0.
CFB = change from Baseline; HAL = Haemophilia Activities List; HJHS = Hemophilia Joint Health Score; PGA = Physician Global Assessment; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; PROMIS = Patient-Reported Outcomes Measurement Information System; SF = Short Form; VAS = visual analog scale.
Note: The PROMIS Pain Intensity 3a psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 PROMIS Pain Intensity 3a data point, and had at least 1 evaluable post-Day 1 PROMIS Pain Intensity 3a data point during the treatment period. a Polyserial correlations are computed with PGA and PGIC scores, and Pearson correlations are computed with the changes in other supportive measures.
Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 4.6.2.10: Ability to Detect Change, Table 26, (pg. 83 of 2016), dated 9 June 2022.

Ability to detect change was further analyzed by the Applicant using mean Haem-A-QoL PH transformed change scores across response categories by change in PGIS Joint, PGIS Activity, PGIC, and change in EQ-5D-5L items at week 26 and week 52 to evaluate whether scores moved in the anticipated direction (e.g., showing greater improvement on mean Haem-A-QoL PH change scores). As shown in Table 13, "between-group differences were statistically significant for almost all comparisons except for PGIS-Activity at Week 26 (α 0.05). Per the Applicant, improvement in mean Haem-A-QoL transformed change scores were seen in subjects who reported improvement on the PGIS-Joint, -Activity, and PGIC at week 52 compared to subjects who reported no change or worsening (effect-size range: -0.51 to -0.43) and statistically significant pairwise tests (α = 0.05) were observed between PGIS response groups at week 52. Moderate to large effect-size estimates of changes (>0.50

in size) were observed between the "1-point improvement" group and the "No Change" or "Worsened" groups of PGIS-Joint and PGIS-Activity."

Table 13. Ability to Detect Change ANOVA Results for Haem-A-QoL PH for pooled sample (Arm A and B), XTEND-1 Study

	Haem-A-QoL PH CFB ANOVA results Mean (SD), median, n	
Responsive subgroup	**Week 26**	**Week 52**
PGIS-Joint		
> 1-point improvement	−18.9 (17.00), −22.5, 14	−28.9 (20.30), −30.0, 14
1-point improvement	−12.1 (16.96), −10.0, 36	−19.7 (23.28), −15.0, 33
No change	−7.8 (18.06), −5.0, 56	−3.2 (14.73), 0.0, 55
Worsened	6.1 (15.46), 5.0, 14	1.4 (12.47), 2.5, 14
ANOVA F-statistic, *P* value; ES_{BL_SD}, ES_{Change_SD}	5.52, *P* = 0.0014; −0.37, −0.47	12.79, *P* < 0.0001; −0.44, −0.51
PGIS-Activity		
> 1-point improvement	−15.4 (18.76), −10.0, 14	−27.7 (22.67), −30.0, 15
1-point improvement	−12.6 (22.22), −10.0, 36	−17.7 (23.87), −15.0, 31
No change	−6.1 (13.94), −5.0, 56	−5.4 (13.75), 0.0, 59
Worsened	−2.9 (20.16), 0.0, 14	6.8 (16.77), 5.0, 11
ANOVA F-statistic, *P* value; ES_{BL_SD}, ES_{Change_SD}	2.09, *P* = 0.1054; −0.37, −0.47	10.66, *P* < 0.0001; −0.44, −0.51
PGIC-Overall		
Very much improved	−8.7 (18.54), −10.0, 23	−21.2 (25.47), −22.5, 26
Much improved	−15.2 (19.17), −10.0, 52	−11.6 (17.99), −10.0, 48
Minimally improved	0.0 (11.67), 0.0, 30	−4.1 (16.96), 0.0, 29
No change	−4.7 (16.98), −5.0, 16	−2.5 (14.85), −2.5, 12
Worsened	22.5 (10.61), 22.5, 2	11.3 (22.87), 12.5, 4
ANOVA F-statistic, *P* value; ES_{BL_SD}, ES_{Change_SD}	5.65, *P* = 0.0003; −0.35, −0.45	4.48, *P* = 0.0021; −0.43, −0.49

ANOVA = analysis of variance; BL = Baseline; CFB = change from Baseline; ES = effect size; Haem-A-QoL = Haemophilia Quality of Life; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; PH = physical health; SD = standard deviation.
Notes: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.
For the overall results, every ESBL_SD value was computed as the overall mean change from non-missing data of the target score and the grouping variable divided by the overall Baseline SD of the target score; every ESChange_SD value was computed as the same mean change divided by the overall SD of change at the same visit with non-missing data of the target score only.
Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.6.2.10: Ability to Detect Change, Table 15. (p. 47 of 2016), dated 9 June 2022.

6.6 Interpretation of Meaningful Within-Patient Score Changes

Three anchor scales were included in the XTEND-1 study: Patient Global Impression of Severity ('Joint symptoms' and 'Physical Activity') and Patient Global Impression of Change (Overall Status). A summary of the frequency of administration of the COAs and the anchor scales can be found in Table 14. Copies of the anchors can be found in Section 7.2.

Patient Global Impression of Severity

The Patient global impression of severity (PGIS) is a single item scale in which patients indicate an overall assessment of their symptoms using a 5-point verbal rating scale, as described in Table 14. Two components of the PGIS were utilized for joint symptoms and physical activity. The PGIS were administered at Baseline, Weeks 26 and 52.

Reviewer's comment(s):

- ***The recall period for the PGIS (i.e., "the past week") did not match the assessment period to calculate the Haem-Q-QoL PH subscale score (i.e., "the past 4-weeks"). This presents a challenge to interpreting anchor-based analyses using the PGIS as an anchor for the PH subscale score, as the PGIS only reflects 1-week out of 4-weeks included in the Haem-A-QoL PH score calculation at baseline, weeks-26, and -52 (i.e., the last week when administered on the same schedule).***
- ***There are concerns with wording and concept alignment of anchors (PGIS Joint) with target COA endpoints ('physical health impairment' for Haem-A-QoL PH). For instance, joint symptoms can encompass multiple concepts (i.e., pain, stiffness, etc.). Anchor scale language was not evaluated during the exit interviews, and thus there are no data to support whether the anchors were interpreted as intended by patients. and aligned well with the Applicant's concept of interest.***

Patient global impression of change (PGIC)

The Patient Global Impression of Change (PGIC) consists of one item adapted to the patient that evaluates all aspects of patients' health using a 7-point verbal rating scale, as described in Table 14. The PGIC was administered at Weeks 26 and 52 to assess if there was an improvement or decline in clinical status since patients started taking the study medication.

Table 14: Summary and Frequency of Assessments of COA Endpoints and Anchor Scales for the XTEND-1 Study.

Endpoint concept/attribute (COA type/name if any)	Anchor (concept)	Anchor response scale	Recall period (target/anchor)	Timing (target/anchor)
Haem-A-QoL Physical Health (PH)	PGIS (Joint symptoms)	5-point Likert: none, mild, moderate, severe, very severe	Past 4-weeks / Past 7-days	Baseline, Weeks 26, 52 (ET) / Baseline, Weeks 26, 52 (ET)
	PGIS (activity)	5-point Likert: none, mild, moderate, severe, very severe	Past 4-weeks / Past 7-days	Baseline, Weeks 26,52 (ET) / Baseline, Weeks 26, 52 (ET)
	PGIC (Overall Status)	7-point VRS: Very much better, moderately better, A little better, The same (No change), A little worse, moderately worse, very much worse	Past 4-weeks / comparison of current state to earlier period ("since you started taking the study medication")	Baseline, Weeks 26,52 (ET)/ Weeks 26, 52
PROMIS Pain Intensity 3a Item 1 (Worst Pain)	PGIS (Joint symptoms)	5-point Likert: none, mild, moderate, severe, very severe	Past 7-days / Past 7-days	Baseline, Weeks 26,52 (ET) / Baseline, Weeks 26, 52 (ET)
	PGIS (activity)	5-point Likert: none, mild, moderate, severe, very severe	Past 7-days / Past 7-days	Baseline, Weeks 26,52 (ET) / Baseline, Weeks 26, 52 (ET)
	PGIC (Overall Status)	7-point VRS: Very much better, moderately better, A little better, The same (No change), A little worse, moderately worse, very much worse	Past 7-days / comparison of current state to earlier period ("since you started taking the study medication")	Baseline, Weeks 26,52 (ET) / Weeks 26, 52

PGIS= Patient Global Impression of Severity; **PGIC** = Patient Global Impression of Change; **ET**= End of Treatment

6.6.1 Anchor-Based Analyses

Haem-A-QoL PH Subscale

The meaningful within-patient change in both PRO-based endpoint scores were evaluated by the Applicant using quantitative anchor-based analyses with the PGIS Joint symptoms (1-point improvement) anchor scale. Figures 4 and 5 show the empirical cumulative distribution function (eCDF) plots of absolute change from baseline in the 4-week Haem-A-QoL PH subscale transformed and raw scores, respectively, at week-52 by PGIS category of change from baseline in Arm A for XTEND-1 study. To facilitate interpretation of the secondary efficacy endpoints based on change from baseline in Haem-A-QoL PH to week 52 in Arm A, the Applicant proposed using a clinically meaningful change threshold range[13] of -6.8 to -4.8, with a point estimate of -6.8 (Table 15).

[13] The range is referred to as meaningful within group change (MWGC) in the Applicant's *PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a.*

Table 15. Results to Construct Clinically Meaningful Change Thresholds for Haem-A-QoL PH, XTEND-1 Study

	Haem-A-QoL PH (CFB to Week 52) 95% CI or 0.2×SD	
Method	**Pooled Arm A and B**	**Arm A**
PGIS-Joint: 1-point improvement	−29.3 to −14.0	−23.2 to −6.8
PGIS-Joint: No change	−6.7 to 1.5	−6.1 to 2.0
PGIS-Activity: 1-point improvement	−26.7 to −9.1	−21.8 to −2.8
PGIS-Activity: No change	−9.3 to −2.2	−8.3 to −1.1
PGIC-Overall: Much improved	−16.8 to −6.3	−12.8 to −2.0
PGIC-Overall: Minimally improved	−10.6 to 2.3	−10.4 to 2.9
0.2 × SD of Baseline	4.7	4.8

CFB = change from Baseline; CI = confidence interval; Haem-A-QoL = Haemophilia Quality of Life; MWGC = Meaningful Within-Group Change; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; PH = physical health; SD = standard deviation.

Note: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.

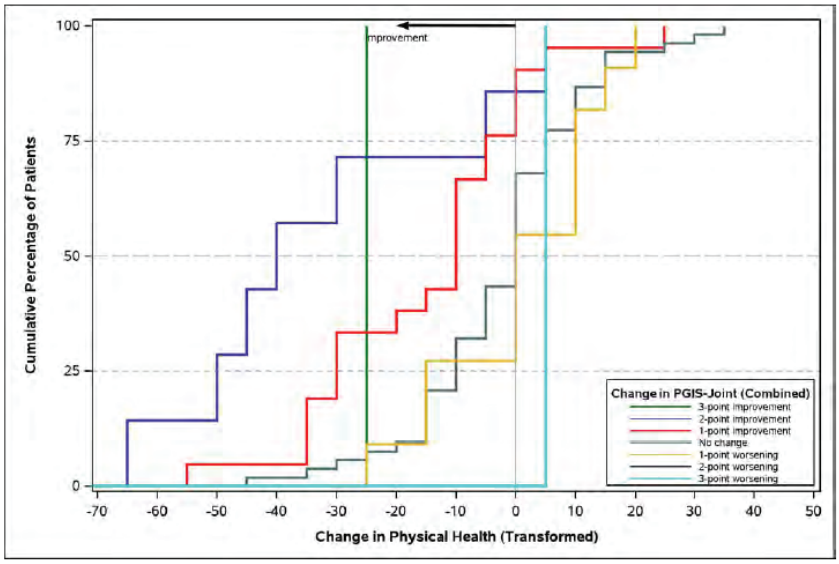
Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 3.7.2: Meaningful Within-Group Change Thresholds, Table 18. (p. 58 of 2016), dated 9 June 2022.

Reviewer's comment(s):

- ***To determine the threshold range of -6.8 to -4.8, the Applicant used the upper bound of the 95% confidence interval (-29.3, -6.8) for the PGIS joint 1-category change to represent the lower range of the meaningful change threshold. The upper range was based on a distribution-based approach (0.2 x SD of Baseline of Haem-A-QoL PH score). Refer to Table 15 above for details. We do not agree with this range because distribution-based methods do not directly take into account the patient voice; and therefore, cannot be used as primary evidence for within-patient clinical meaningfulness[14].***
- ***Applicant's original anchor-based analyses used the 0-100 transformed scores for the Haem-A-QoL PH. To aid in interpretation of the endpoint results, the Agency issued an information request (dated 16 November 2022) and requested additional anchor-based analyses using raw scores (e.g., eCDF and probability density function [PDF] curves). Specifically, the Applicant provided new eCDF (Figure 5) and PDF (Figure 6) curves using the raw scores on the Haem-A-QoL PH. For further details on the Applicant's anchor-based analyses using the transformed scores, refer to Applicant's PRO Evidence Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a [dated 9 June 2022].***

[14] Patient-Focused Drug Development Guidance Workshop: Incorporating Clinical Outcome Assessment into Endpoints for Regulatory Decision-Making. Section C. Other Methods, (p. 22); dated 6 December 2019. https://www.fda.gov/media/132505/download

Figure 4. eCDF Plot of Change from baseline in Haem-A-QoL PH Transformed Scores to Week 52 by PGIS Joint Categories of Change from Baseline (Arm A, XTEND-1 Study)



Change in PGIS-Joint (Combined)	n	P10	P25	Median	P75	P90
3-point improvement	1	−25.0	−25.0	−25.0	−25.0	−25.0
2-point improvement	7	−65.0	−50.0	−40.0	−5.0	5.0
1-point improvement	21	−35.0	−30.0	−10.0	−5.0	0.0
No change	53	−15.0	−10.0	0.0	5.0	15.0
1-point worsening	11	−15.0	−15.0	0.0	10.0	15.0
2-point worsening	1	5.0	5.0	5.0	5.0	5.0
3-point worsening	1	5.0	5.0	5.0	5.0	5.0

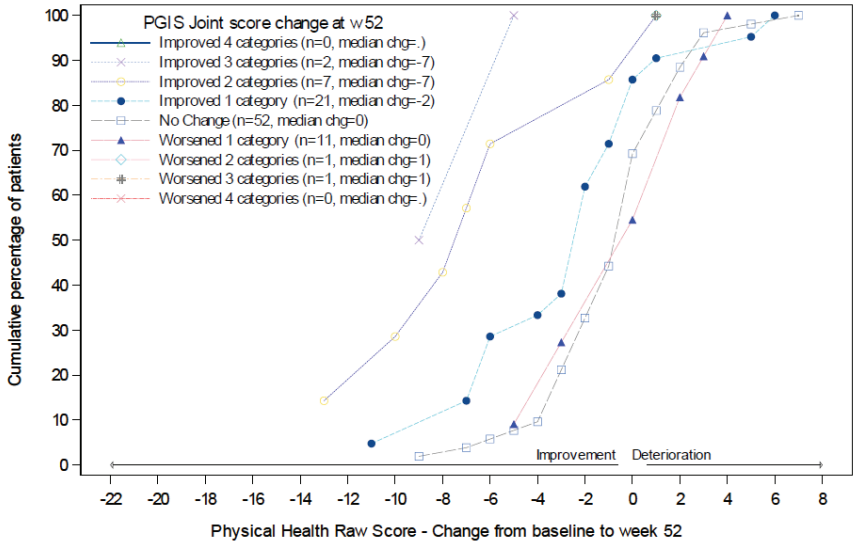
CFB = change from Baseline; eCDF = empirical cumulative distribution function; Haem-A-QoL = Haemophilia Quality of Life; P = percentile; PGIS = Patient Global Impression of Severity; PH = physical health.

Notes: The Haem-A-QoL psychometric analysis sample was defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data point, and had at least 1 evaluable post-Day 1 Haem-A-QoL data point during the treatment period.

The very severe and severe subgroups were combined before the computation of PGIS change.

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a , Section 3.7.1.2:Estimates from the XTEND-1 Study, Figure 4. (pg. 53 of 2016), dated 9 June 2022.

Figure 5. eCDF Plot of Change from baseline in Haem-A-QoL PH Raw Scores to Week 52 by PGIS Joint Categories of Change from Baseline (Arm A, XTEND-1 Study)

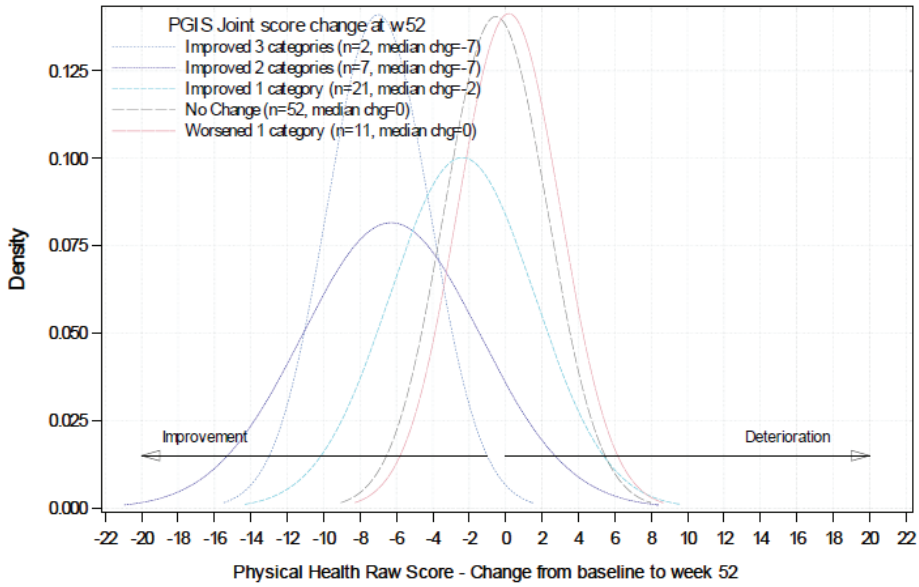


The Haem-A-QoL psychometric analysis sample was defined as all participants >=17 years old who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data, and had at least 1 evaluable post-Day 1 Haem-A-QoL data during the treatment period.

Median change from baseline in Haem-A-Qol raw score are provided for each group of anchor

Source: Response to FDA Information Request, Figure 5 (p. 22 of 42, dated 02 December 2022.

Figure 6. Probability Density Functions of Change from Baseline in Haem-A-QoL PH raw score at week 52 according to change from baseline in PGIS Joint at week 52- Arm A, XTEND-1 Study



The Haem-A-QoL psychometric analysis sample was defined as all participants >=17 years old who received at least 1 dose of the study intervention, had an evaluable Day 1 Haem-A-QoL data, and had at least 1 evaluable post-Day 1 Haem-A-QoL data during the treatment period.

Median change from baseline in Haem-A-QoL raw score are provided for each group of the anchor.

Source: Response to FDA Information Request, Figure 11 (p. 28 of 42); dated 02 December 2022.

Reviewer's comment(s):
Given the study design, small sample size in Arm A, and the even smaller sample size for each anchor change category, the eCDF and PDF plots should be interpreted with caution. It is also worth noting that the majority of patients reported "no change" on the PGIS-Joint (n=52). When visually examining the eCDF curves in Figure 5 (also illustrated by the PDF curves in Figure 6), the separation between "improved 1 category" and "no change" curves was minimal; and there is no separation between the "no change" and "worsened 1 category" curves. In addition, the minimal change in raw scores (Table 4) makes it difficult to assess whether there truly is a clinically meaningful within patient change that reflects improvement, whether or not the purported change can be attributed to treatment given the study design.

PROMIS Pain Intensity 3a Item (Worst Pain)

To facilitate interpretation of the PROMIS Pain Intensity item (Worst Pain) from change in baseline to week 52, the Applicant's proposed meaningful change threshold range[15] was -0.6 to -0.2, with a point estimate of -0.6. Per the sponsor, *"For the PROMIS Pain Intensity 3a Worst Pain, the upper limit of the 95% CI for the 1-point improvement in PGIS-Joint subgroup in Arm A using all patients (aged ≥ 12 years) was −0.6; the lower limit of the 95% CI for the no change in PGIS-Joint subgroup was −0.3 Table 16). The 0.2 × SD values was ±0.2. Therefore, the proposed MWGC threshold range for the PROMIS Pain Intensity 3a Worst Pain is −0.6 to −0.2, with a point estimate of −0.6, based on all patients (aged ≥ 12 years)." (PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a, Section 4.7.3 (p. 97 of 2016), received date 30 June 2022).*

Table 16. Results to Construct the Range of MWGC Thresholds for PROMIS Pain Intensity 3a Worse Pain Score, Pooled Adults and Adolescents, XTEND-1 Study

	PROMIS Pain Intensity 3a Worst Pain (CFB to Week 52) 95% CI or 0.2×SD	
Sample/method	Pooled Arm A and B	Arm A
PGIS−Joint: 1−point improvement	−1.4 to −0.8	−1.5 to −0.6
PGIS−Joint: No change	−0.3 to 0.2	−0.3 to 0.2
0.2 × SD of Baseline	0.2	0.2

CFB = change from Baseline; CI = confidence interval; MWGC = Meaningful Within-Group Change; PGIS = Patient Global Impression of Severity; PROMIS = Patient-Reported Outcomes Measurements Information System; SD = standard deviation.
Note: The PROMIS Pain Interference 3a psychometric analysis sample as defined as all participants who received at least 1 dose of the study intervention, had an evaluable Day 1 PROMIS Pain Interference 3a data point, and had at least 1 evaluable post-Day 1 data point during the treatment period.

[15] The range is referred to as meaningful within group change (MWGC) in the Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a.

Source: Applicant's PRO Dossier for the Haem-A-QoL and PROMIS Pain Intensity 3a , Section 4.7.2: Meaningful Within-Group Change Thresholds, Table 31. (pg. 96 of 2016), dated 9 June 2022.

Reviewer's comment(s):

- ***To determine the threshold range of -0.6 to -0.2, the Applicant used the upper bounds of the 95% confidence interval (-1.5, -0.6) for the PGIS joint 1 category change to represent the lower range of the meaningful change threshold. The upper range was calculated using a distribution-based formula created by the Applicant (0.2 x SD of Baseline of Haem-A-QoL PH score). Refer to Table 16 above for values. We do not agree with this range because distribution-based methods do not directly take into account the patient voice; and therefore, cannot be used as primary evidence for within-patient clinical meaningfulness[16].***

To further evaluate the PROMIS Pain Intensity 3a, the Agency examined the individual items as reference measures and found that the single item (worst pain) was simple and appeared to be well understood by the optional exit interview participants (refer to Section 6.4). Therefore, to aid in interpretation of the endpoint results, the Agency issued an Information Request (SDN 20, received dated 02 December 2022) asking for additional analyses in which the single-item raw scores (worst pain) were plotted against the primary endpoint (ABR) to evaluate whether pain reduction correlated with the anticipated treatment effect at the study timepoints. Figure 7 depicts bubble plots demonstrating the change in PROMIS Pain Intensity 3a Item (Worst Pain) versus the primary endpoint, annualized bleeding rate (ABR) at baseline, week 26, and week 52. As shown in Figure 7, very weak correlations between the PROMIS Pain Intensity 3a (Worst Pain) item and ABR were observed at baseline (Spearman Correlation = 0.13), and at weeks -26 and -27 (Spearman Correlation = 0.17), respectively.

Reviewer's comment(s):
A very weak correlation ≤ 0.17) between the PROMIS Pain Intensity Item (Worst Pain) and ABR suggests that pain intensity at its worst may not be responsive to treatment, or that the measure may be insensitive to detect change. As such, it would be misleading to suggest that prophylactic treatment leads to an 'improvement' in pain intensity for the indicated study population.

[16] Patient-Focused Drug Development Guidance Workshop: Incorporating Clinical Outcome Assessment into Endpoints for Regulatory Decision-Making. Section C. Other Methods, (p. 22); dated 6 December 2019. https://www.fda.gov/media/132505/download

Figure 7. Bubble plots of PROMIS Pain Intensity 3a Pain at is Worst Score versus Annualized Bleeding Rate (ABR) at (from left to right, and down) Baseline, Week 26, and Week 52 for XTEND-1 Arm A.

Figure 1 - Bubble plots of PROMIS Pain Intensity 3a - Pain at its worst score at baseline versus annualized bleeding rate (ABR) - Arm A - Full Analysis Set

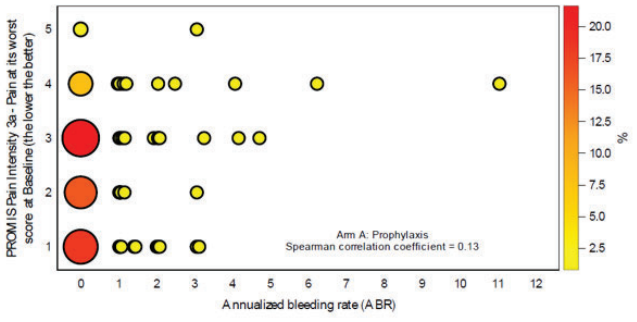


Figure 2 - Bubble plots of PROMIS Pain Intensity 3a - Pain at its worst score at Week 26 versus annualized bleeding rate (ABR) - Arm A - Full Analysis Set

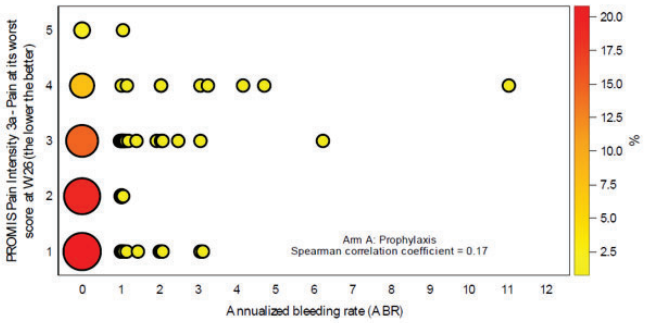
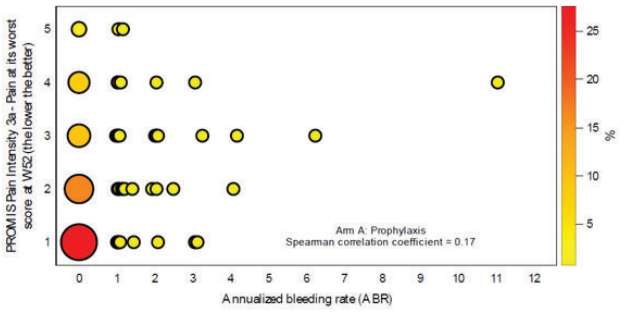


Figure 3 - Bubble plots of PROMIS Pain Intensity 3a - Pain at its worst score at Week 52 versus annualized bleeding rate (ABR) - Arm A - Full Analysis Set



Source: Response to FDA Information Request, Figures 1-3, (p. 6-8 of 42); dated 02 December 2022.

Summary of analyses

- The Applicant provided anchor-based, eCDF, and PDF curves for each of the three anchors (PGIS Joint and Physical Activity; and PGIC Status) for each of the PRO measures.
- The Applicant provided bubble plots to demonstrate the correlation of change from baseline in Pain Intensity 3a (Worst Pain) with annualized bleeding rate (ABR) at baseline, week 26, and week 52.
- Qualitative data included optional exit interviews were conducted to help interpret what constituted a clinically meaningful within patient change for both the Haem-A-QoL and PROMIS Pain Intensity 3a Item (Worst Pain). For results, refer to Section 6.4.

Reviewer's comment(s):

- ***The recruitment strategy (i.e., convenience sampling) for the optional exit interviews may have led to biased responses. The exit interview participants who opted to participate may have included a subset of patients from the XTEND-1 trial who observed a larger magnitude of change in physical function from baseline to week-52. As a result, the bias due to selection would be expected to be reflected during the interviews. For instance, 82.8% (n=24/29) of exit interview patients reported being 'very satisfied' with the study treatment and all (n=29/29, 100%) said they preferred***

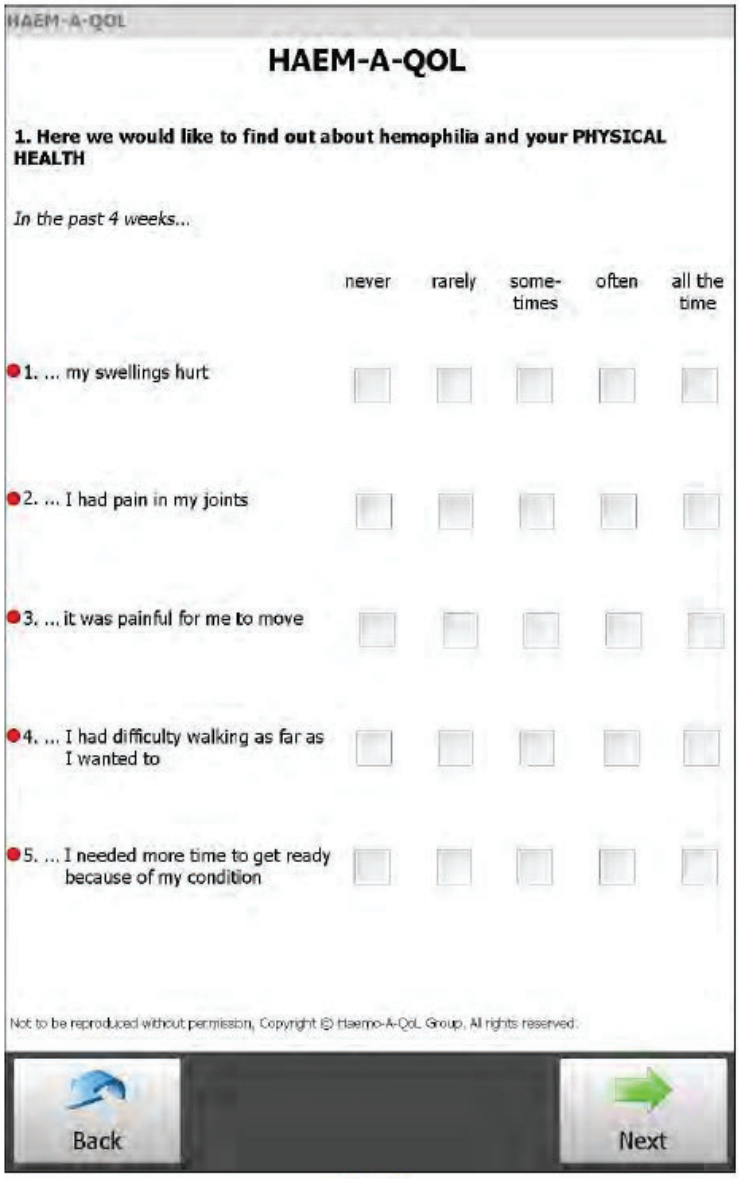
the treatment over their previous hemophilia treatment. In turn, all participants (n=27/27) responded that any positive change (i.e., 1 or more) would be meaningful to them with a 2-point change being even better on the PROMIS Pain Intensity 3a measure. The proportion and sample size of those reporting the magnitude of change that was meaningful on the Haem-A-QoL PH was not provided. The sponsor reported "…participants consistently reported that any improvement in the concepts assessed by the Haem-A-QoL PH subscale would be meaningful to them; 1 participant even indicated that remaining at the same level would be meaningful since that would suggest that further damage to joints was not occurring." (PRO Evidence Dossier for the Haem-A-QoL PH and PROMIS Pain Intensity 3a, p. 31 of 2016, received date 30 June 2022).

- ***The exit interview guide did not include questions about stability, or staying the same/no change, and whether those outcomes were meaningful to patients. The questions included in the guide focused on improvement for both the Haem-A-QoL and PROMIS Pain intensity 3a item (worst pain). The psychometric and anchor-based analysis data from both PRO measures suggests that 'staying the same' or 'no change' may be just as meaningful as a 1 category change to patients.***
- ***Of the 28 exit interview participants who reported pretrial issues with physical health/function, 26 reported improvements in at least 1 of the Haem-A-QoL PH concepts (at least 75% of the interviewees for each concept). However, this is inconsistent with the data from the XTEND-1 Arm A results, in which a large proportion of participants from Arm A for the XTEND-1 study reported 'none' to 'moderate' (>50%) for items 1 and 3-5 on the Haem-A-QoL PH at baseline and provided similar responses at 52-weeks using raw scores, indicating a potential floor effect.***
- ***There are similar concerns regarding potential floor/ceiling effects with the PROMIS Pain Intensity 3a item. Twenty-eight exit interview participants (across both Arms A and B) reported pretrial hemophilia-related pain as at least of moderate or severe intensity on the PROMIS instrument; of these, 25 reported improvements in pain during the XTEND-1 trial during their interview. However, a large proportion of patients in the XTEND-1 study (Arm A) reported low raw scores for the PROMIS Pain Intensity 3a item (worst pain) at baseline (~2.47) and their pain ratings were largely stable over the 52-weeks suggesting potential floor effects which may be supported by the optional exit interviews with due caution given the convenience sampling.***

6. APPENDICES

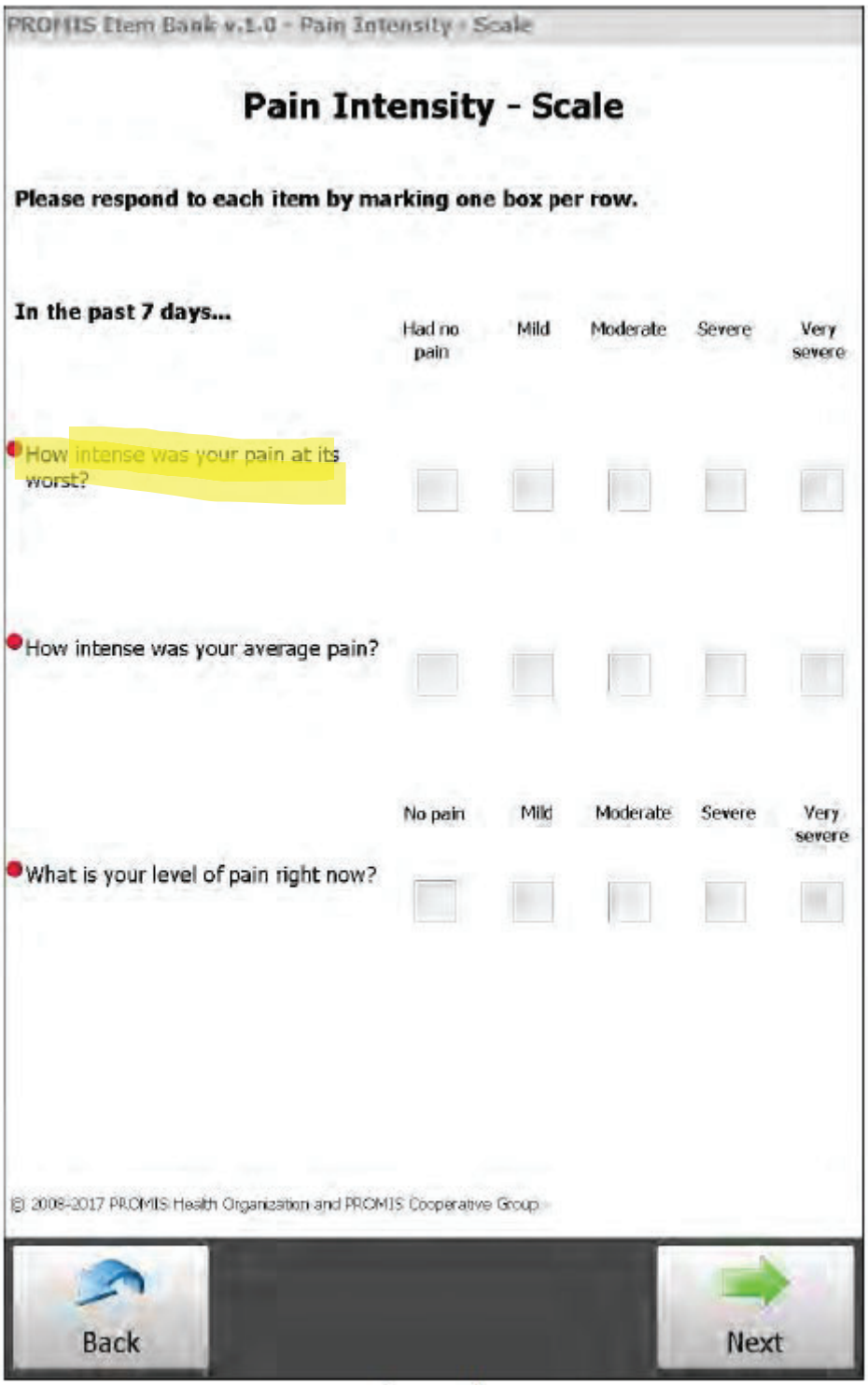
7.1 Appendix A: Copy of COAs

7.1.1 Haem-A-QoL PH



Screen 1

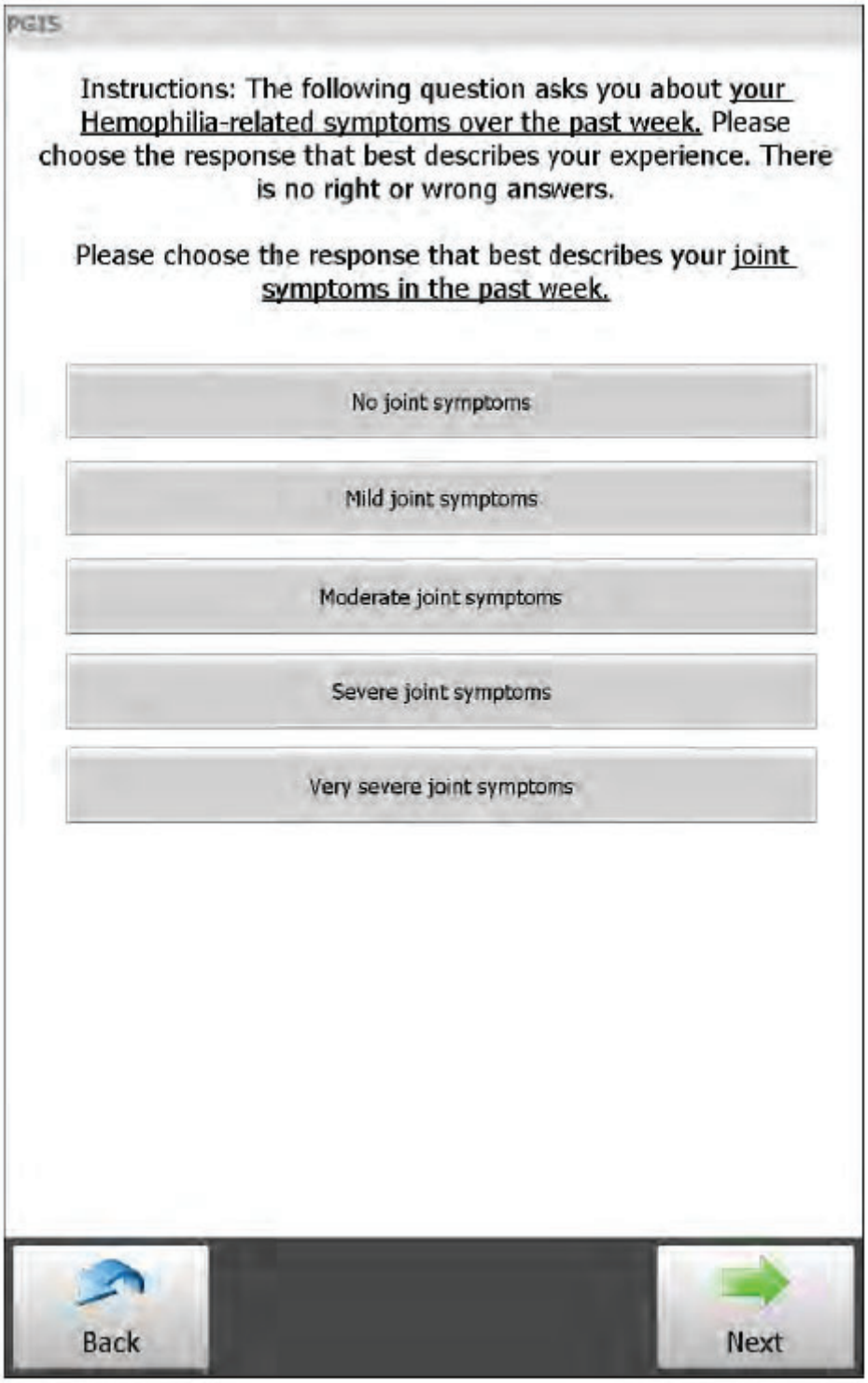
7.1.2 PROMIS Scale v1.0 Pain Intensity 3a



Screen 1

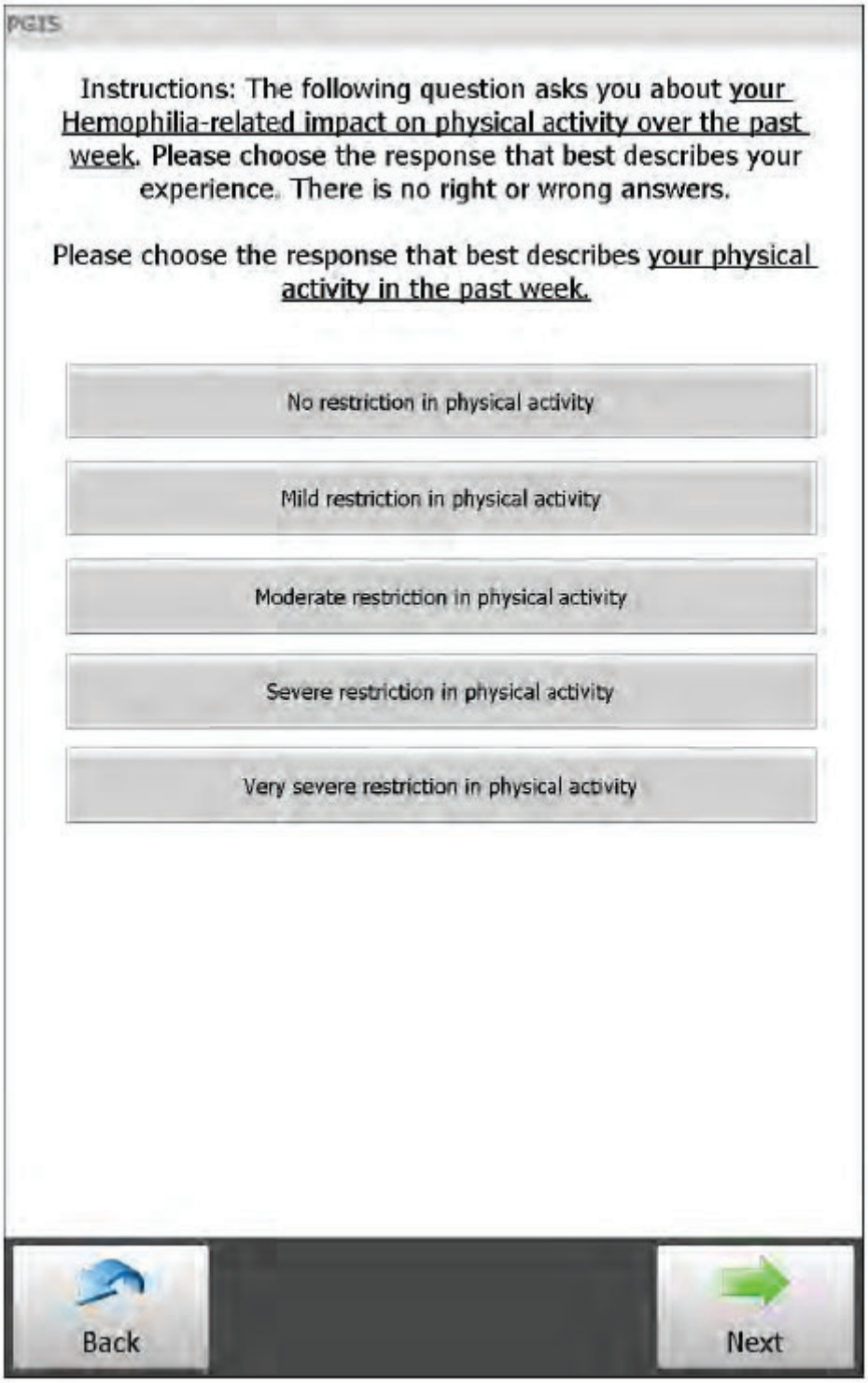
7.2 Appendix B: Copy of Anchor Scales

7.2.1 PGI-S Joint Symptoms



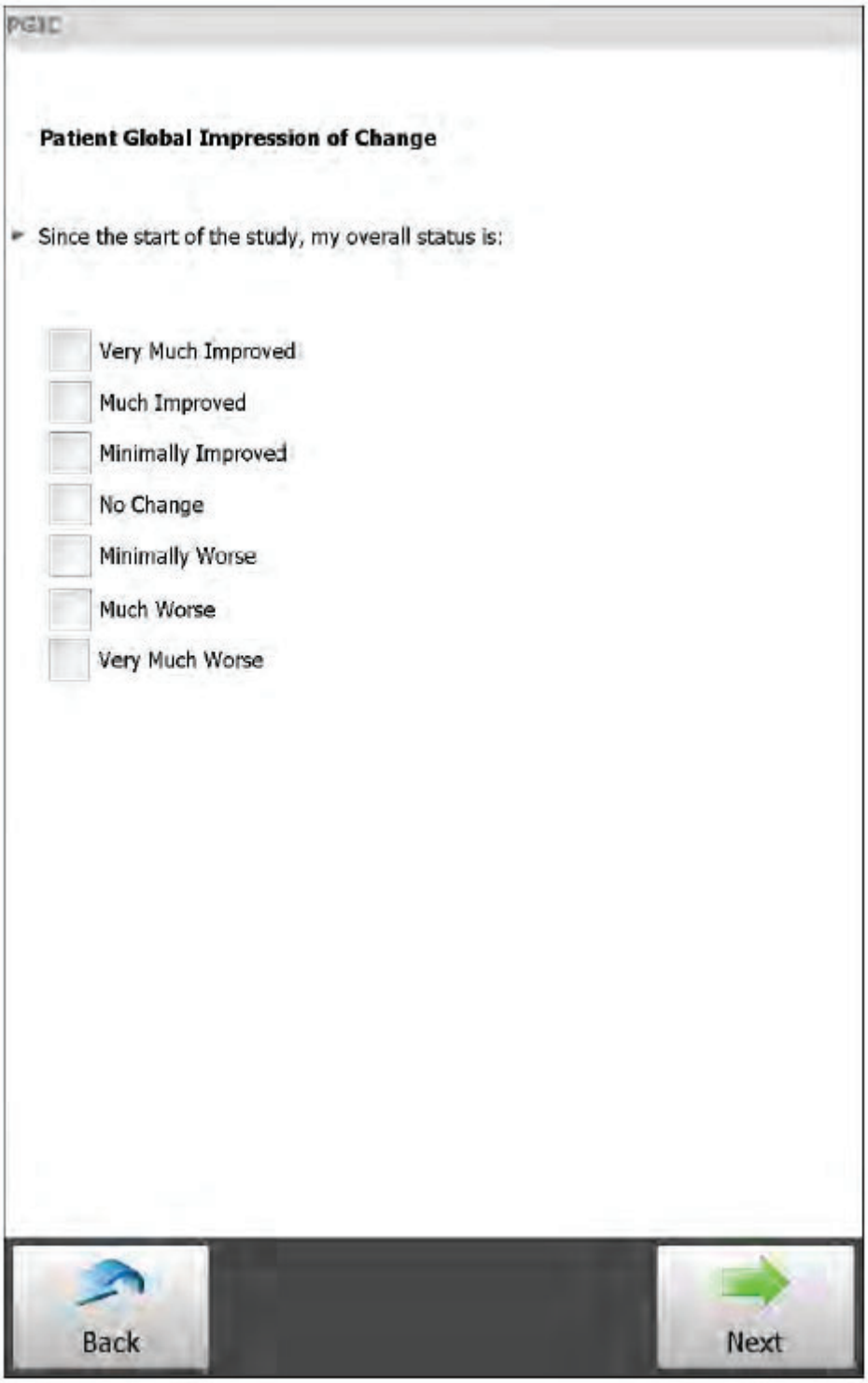
Screen 2

7.2.2 PGI-S Physical Activity



Screen 3

7.2.3 PGI-C Overall Status



Screen 2

7.3 Materials Reviewed

Document	SDN	eCTD#	Date Received
1.14.1.2 Altuviio-Initial BLA- AnnotatedPI-PPI-JUN-2022 (MS Word)	*1*	*0002*	30Jun2022
16.2.6 EFC16293 16.2.6 Efficacy response data	1	0002	30Jun2022
16.1.1 EFC16293 16.1.1 Protocol	1	0002	30Jun2022
5.3.5.1 EFC16293 Synopsis	1	0002	30Jun2022
16.2.4 EFC16293 Demographic data, data at baseline and medication details	1	0002	30Jun2022
5.3.5.1 Clinical Outcome Assessment Dossier	1	0002	30Jun2022
2.5 Clinical Overview	1	0002	30Jun2022
5.3.5.1 EFC16293 1 to 8 Study Report Body	1	0002	30Jun2022
1.11.3 Response to FDA Request dated 16-Nov-2022 – Clinical/DCOA	20	0021	02Dec2022
Publications			
FDA. Food and Drug Administration. The voice of the patient: a series of reports from the U.S. Food and Drug Administration's (FDA's) patient-focused drug development initiative hemophilia A, hemophilia B, von Willebrand disease and other heritable bleeding disorders. 2016. https://www.fda.gov/files/about%20fda/published/The-Voice-of-the-Patient--Hemophilia-A--Hemophilia-B--von-Willebrand-Disease-and-Other-Heritable-Bleeding-Disorders.pdf. *von Mackensen S, Gringeri A. Quality of life in hemophilia. In: Preedy VR, Watson RR, editors. Handbook of Disease Burdens and Quality of Life Measures. New York, NY: Springer; 2010.* *von Mackensen S, Gringeri A, Haem-A-QoL study Group. Health-related quality of life in adult patients with haemophilia – Assessment with a new disease-specific questionnaire (Haem-A-QoL). J Thrombo Haem. 2005a;3(Suppl 1):P0813.* *Skinner MW, Negrier C, Paz-Priel I, Chebon S, Jimenez-Yuste V, Callaghan MU, et al. The effect of emicizumab prophylaxis on long-term, self-reported physical health in persons with haemophilia A without factor VIII inhibitors in the HAVEN 3 and HAVEN 4 studies. Haemophilia. 2021 Sep;27(5):854-65. doi:http://dx.doi.org/10.1111/hae.14363.* *Paredes AC, Teixeira P, Almeida A, Pinto PR. Prevalence and interference of chronic pain among people with Hemophilia: a systematic review and meta-analysis. J Pain. 2021 Oct;22(10):1134-45. doi:http://dx.doi.org/10.1016/j.jpain.2021.03.157.*			